Universidade de Aveiro Departamento de Matemática
2005

Rute Correia Lemos

Contradomínio Numérico, Desigualdades Matriciais e suas Aplicações em Física

Universidade de Aveiro Departamento de Matemática
2005

Rute Correia Lemos

Contradomínio Numérico, Desigualdades Matriciais e suas Aplicações em Física

Dissertação apresentada à Universidade de Aveiro, para cumprimento dos requisitos necessários à obtenção do grau de Doutor em Matemática, realizada sob a orientação científica da Prof. Doutora Natália Bebiano, Professora Catedrática do Departamento de Matemática, Faculdade de Ciências e Tecnologia da Universidade de Coimbra.

Universidade de Aveiro Departamento de Matemática
2005

Rute Correia Lemos

Numerical Range, Matrix Inequalities and its Applications in Physics

Thesis presented to the University of Aveiro, in partial fulfilment of the requirements for the degree Doctor of Philosophy in Mathematics, under the scientific supervision of Professor Natália Bebiano, Full Professor at the Department of Mathematics, Faculty of Sciences and Technology of the University of Coimbra.

júri

presidente	Doutor Dinis Gomes Magalhães dos Santos Professor Catedrático da Universidade de Aveiro, em representação da Ex.ma Reitora da Universidade de Aveiro
vogais	Doutora Natália Isabel Quadros Bebiano Pinheiro da Providência e Costa Professora Catedrática da Faculdade de Ciências e Tecnologia da Universidade de Coimbra (Orientadora)
	Doutora Maria Paula Macedo Rocha Malonek Professora Catedrática da Universidade de Aveiro
	Doutora Maria Graça Nunes da Silva Rendeiro Marques Professora Associada da Faculdade de Ciências e Tecnologia da Universidade do Algarve
	Doutor João Carlos David Vieira Professor Associado Aposentado da Universidade de Aveiro
	Doutora Susana Margarida Figueiredo de Sousa Borges Furtado Professora Auxiliar da Faculdade de Economia da Universidade do Porto
	Doutora Rosa Amélia Batista Ferreira Soares Martins Professora Auxiliar da Universidade de Aveiro

agradecimentos

Sendo o trabalho científico, mais do que um mero esforço pessoal, um esforço colectivo, muito tenho a agradecer de forma reconhecidamente especial:

- à Prof. Doutora Natália Bebiano, pela preciosa orientação e pronto acompanhamento, pelo seu exemplo de excelência, rigor, perseverança e entusiasmo, pela sua cultura científica e atitude face à actividade de investigação, que me ensinaram a traçar, com paixão, a minha rota, neste mar infinito da Matemática. Mais do que um prazer, foi um privilégio;
- ao Prof. Doutor João da Providência, pela enorme disponibilidade, pelas ideias perspicazes e sugestões que motivaram, em particular, a ligação que neste trabalho se faz entre a Matemática e a Física;
- à Prof. Doutora Graça Soares, pelas discussões, viagens e projectos que partilhámos;
- ao Departamento de Matemática e Unidade de Investigação e Desenvolvimento "Matemática e Aplicações" da Universidade de Aveiro, pelas condições que me proporcionaram;
- ao Departamento de Matemática da Universidade de Coimbra, pelo amável acolhimento;
- aos meus pais, pelo apoio que sempre me prestaram e constante incentivo;
- ao David, por todo o carinho, dedicação e tolerância;
- à minha família;
- aos meus amigos.

resumo

Esta dissertação abrange o estudo de contradomínios numéricos, de desigualdades matriciais e suas incidências em Física, do ponto de vista da álgebra linear e da teoria de operadores lineares.

Discutem-se propriedades geométricas do contradomínio numérico de operadores num espaço com produto interno indefinido. Estendem-se alguns contradomínios de aplicações complexas do grupo unitário, envolvendo o traço ou o determinante, ao grupo pseudo-unitário. Estudam-se a curva geradora de fronteira, algumas formas especiais e os pontos angulosos destes conjuntos, generalizando resultados clássicos conhecidos. Em particular, apresentam-se classes de matrizes cujas curvas algébricas geradoras de fronteira são hipérboles.

Em sistemas físicos compostos de várias partículas idênticas, é útil definir operadores que criam ou destroem uma partícula num estado individual específico. Em Física Quântica, definem-se operadores de emparelhamento em termos daqueles operadores. Descrevem-se algumas propriedades espectrais destes operadores de emparelhamento e caracterizam-se os respectivos contradomínios numéricos. No contexto da teoria de matrizes, os resultados originam os contradomínios numéricos de certas matrizes tridiagonais finitas e infinitas, que se prova terem a forma elíptica ou hiperbólica.

Finalmente, deduzem-se algumas desigualdades espectrais de matrizes, envolvendo a noção de entropia relativa quântica. Generaliza-se a desigualdade termodinâmica e prova-se uma cadeia de afirmações equivalentes, envolvendo esta desigualdade e a desigualdade de Peierls-Bogoliubov.

Palavras-chave: contradomínio numérico, espaço com produto interno indefinido, operadores de criação e destruição, entropia relativa, desigualdades matriciais.

abstract

This dissertation is devoted to the study of numerical ranges, matrix inequalities and their incidence in Physics, from the point of view of linear algebra and linear operator theory.

Geometric properties of numerical ranges of operators on an indefinite inner product space are discussed. Some ranges of complex valued functions in the unitary group, involving the trace or the determinant, are extended to the pseudo-unitary group. Boundary generating curves, special shapes and corners of these sets are studied, which generalize some classical known results. In particular, classes of matrices are presented such that the boundary generating algebraic curves are hyperbolas.

For physical systems composed of many identical particles, it is useful to define operators that create or annihilate a particle in a specified individual state. In quantum physics, pairing operators are defined in terms of those operators. Spectral properties of pairing operators are described and their numerical ranges are characterized. In the context of matrix theory, the results provide the numerical ranges of certain finite and infinite tridiagonal matrices, which are proved to have either the elliptical or the hyperbolical shape.

Finally, some matrix spectral inequalities, involving the notion of quantum relative entropy, are derived. The thermodynamic inequality is generalized and a chain of equivalent statements, involving this inequality and the Peierls-Bogoliubov inequality, is proved.

Keywords: numerical range, indefinite inner product space, creation and annihilation operators, relative entropy, matrix inequalities.

	União Europeia Fundo Social Europeu	

Índice

Notação

$\mathbb{N}$	conjunto dos números naturais
$\mathbb{N}_0$	conjunto dos números inteiros não negativos
$\mathbb{Z}$	conjunto dos números inteiros
$\mathbb{R}$	corpo dos números reais
$\mathbb{C}$	corpo dos números complexos
$\mathcal{H}$	espaço de Hilbert complexo separável
$\mathbb{R}^n$	espaço vectorial real dos n-uplos reais
$\mathbb{C}^n$	espaço vectorial dos n-uplos complexos
M_n	álgebra das matrizes $n \times n$ com entradas complexas
$M_{n,m}$	espaço vectorial das matrizes $n \times m$ com entradas complexas
$M_n(\mathbb{R})$	álgebra das matrizes $n \times n$ com entradas reais
H_n	espaço vectorial real das matrizes Hermíticas de ordem n
$U_{r,n-r}$	grupo das matrizes pseudo-unitárias de assinatura $(r, n-r)$
S_m	grupo simétrico de grau m de identidade id
$\otimes^m \mathcal{H}$	m-ésimo produto tensorial de $\mathcal{H}$
$\bigwedge^m \mathcal{H}$	m-ésimo espaço exterior, de Grassmann ou anti-simétrico sobre $\mathcal{H}$
$\mathcal{H}_{(m)}$	m-ésimo espaço (completamente) simétrico sobre $\mathcal{H}$
$\oplus$	soma directa
$\otimes$	produto tensorial
$\wedge$	produto anti-simétrico
$*$	produto simétrico
$\overline{K}$	fecho topológico do conjunto K
∂K	fronteira do conjunto K
$\operatorname{co} K$	invólucro convexo do conjunto K
$\operatorname{Re} K$	projecção ortogonal do subconjunto K do plano no eixo real
$\operatorname{Im} K$	projecção ortogonal do subconjunto K do plano no eixo imaginário
$\operatorname{Re} z$	parte real do número complexo z
$\operatorname{Im} z$	parte imaginária do número complexo z
$\operatorname{diag}(a_1, \ldots, a_n)$	matriz diagonal de entradas principais os escalares $a_1, \ldots, a_n$

I_n	matriz identidade de ordem n
J	$I_r \oplus -I_{n-r}$, $0 \leq r \leq n$
x^*	vector linha de componentes conjugadas do vector coluna $x \in \mathbb{C}^n$
A^*	adjunta ou transconjugada de A
$A^{[*]}$	adjunta-H ou adjunta-J de A
A^{-1}	inversa da matriz não-singular A
$A^{1/2}$	única raiz quadrada semi-definida positiva da matriz A
$A^{(k)}$	k-ésima matriz composta da matriz A
$A[kl]$	submatriz principal de A determinada pelas linhas e colunas k e l
$P_m^{(r)}A$	r-ésima derivação da matriz ou operador A em $\mathbb{C}^n_{(m)}$ ou $\mathcal{H}_{(m)}$
$A > 0$	$A \in H_n$ definida positiva
$A \geq 0$	$A \in H_n$ semi-definida positiva
$A \geq B$	$A, B \in H_n$, satisfazendo $A - B \geq 0$
$\sigma(A)$	espectro do operador ou matriz A
$\mathrm{Tr}\,(A)$	traço da matriz A
$\det\,(A)$	determinante da matriz A
$\mathrm{per}\,(A)$	permanente da matriz A
$\mathrm{Tr}_m^{(k)}(A)$	traço parcial da matriz A
$\Gamma_{m,n}$	conjunto das sucessões de m números inteiros de 1 a n
$Q_{m,n}$	subconjunto de $\Gamma_{m,n}$ das sucessões estritamente crescentes
$G_{m,n}$	subconjunto de $\Gamma_{m,n}$ das sucessões não decrescentes
$e_\alpha^\otimes$	$e_{\alpha_1} \otimes \cdots \otimes e_{\alpha_m}$, em que $\alpha = (\alpha_1, \ldots, \alpha_m) \in \Gamma_{m,n}$
$e_\alpha^\wedge$	$e_{\alpha_1} \wedge \cdots \wedge e_{\alpha_m}$, em que $\alpha = (\alpha_1, \ldots, \alpha_m) \in Q_{m,n}$
e_α^*	$e_{\alpha_1} * \cdots * e_{\alpha_m}$, em que $\alpha = (\alpha_1, \ldots, \alpha_m) \in G_{m,n}$
e_i^k	$e_i * \cdots * e_i$, envolvendo k factores
δ_{ij}	símbolo de Kronecker, isto é, 1, se $i = j$, e 0, se $i \neq j$
Γ^*	álgebra simétrica sobre $\mathbb{C}^2$
$\Gamma^\wedge$	álgebra de Grassmann sobre $\mathbb{C}^2$
$\Gamma^{(q)}$	subespaço de Γ^* gerado pelos vectores $e_1^{n+\tau_q} * e_2^{n+\kappa_q}$, $n \in \mathbb{N}_0$
$S(A)$	entropia da matriz semi-definida positiva A
$S(A, B)$	entropia relativa de Umegaki
$\hat{S}(A\|B)$	entropia relativa para operadores (de Fujii e Kamei)
$\#_\alpha$	média da potência-α de matrizes, $0 \leq \alpha \leq 1$
$\prec$	majoração de matrizes Hermíticas
$\prec_w$	majoração fraca de matrizes Hermíticas
$\prec_{(\log)}$	majoração logarítmica de matrizes semi-definidas positivas

Introdução

In early studies of Hilbert spaces (by Hilbert, Hellinger, Toeplitz, and others) the objects of chief interest were quadratic forms.

Paul R. Halmos, in *A Hilbert Space Problem Book*

A noção de contradomínio numérico deve parte da sua motivação à teoria clássica de formas quadráticas. Seja $\mathcal{H}$ um espaço de Hilbert complexo separável dotado de um produto interno $\langle \cdot, \cdot \rangle$ e seja

$$x \;\mapsto\; \langle Ax, x \rangle, \qquad x \in \mathcal{H}, \tag{1}$$

a forma quadrática associada ao operador linear A em $\mathcal{H}$. Considera-se $\mathcal{H}$ munido da norma induzida pelo produto interno $\|x\| = \langle x, x \rangle^{1/2}$, $x \in \mathcal{H}$. O contradomínio da forma quadrática (1) restrita à superfície da esfera unitária $\{x \in \mathcal{H} : \|x\| = 1\}$ denomina-se o *contradomínio numérico* de A e denota-se por $W(A)$. A literatura sobre este conjunto remonta ao ano de 1918, nomeadamente, a dois artigos famosos, um de Toeplitz [127] e outro de Hausdorff [75]. Em conjunto, estabelecem uma das propriedades geométricas mais importantes do contradomínio numérico: a sua convexidade. Desde então, este resultado, conhecido por *Teorema de Toeplitz-Hausdorff* tem suscitado o interesse de uma vasta comunidade de investigadores, sendo considerável a variedade de generalizações ao conceito clássico, com motivações teóricas e aplicadas. Particularmente interessante é o estudo da relação entre as propriedades algébricas e analíticas do operador e as propriedades geométricas do seu contradomínio numérico.

A teoria de contradomínios numéricos e suas generalizações tem-se revelado de sublinhável alcance em aplicações a diferentes ramos da matemática pura e aplicada, como a análise matricial, teoria de operadores, análise funcional, análise numérica, álgebras de Banach, álgebras-C^*, teoria de sistemas, equações diferenciais, tendo-se estendido recentemente à Física [15, 17, 18, 117] e à Arquitectura [109]. A riqueza desta teoria advém não só do largo espectro das suas aplicações, mas também se reflecte na utilização de métodos e técnicas de vários domínios da matemática, tais como algébra linear e multilinear, análise, teoria combinatória, teoria das perturbações, grupos e álgebras de Lie, geometria diferencial, curvas algébricas, topologia algébrica e diferencial, entre outros.

O contradomínio numérico tem interesse em Física por várias razões, já que encerra informação sobre o operador A, particularmente, sobre a localização dos seus valores próprios. Se $x \in \mathcal{H}$ representar um estado (normalizado) de um sistema com uma só partícula, e se o comportamento dessa partícula for determinado pelo observável representado pelo operador A, então $\langle Ax, x \rangle$ é o valor médio de uma medida do observável A no estado x. O conjunto $W(A)$ pode ser entendido como a colecção de todos os valores médios possíveis de medidas de A em estados do sistema em causa [17]. O comportamento da fronteira de $W(A)$ pode conter informação particularmente útil sob o ponto de vista físico. Contudo, esta terminologia de contradomínios numéricos está para já ausente dos textos de Física.

A substituição do produto interno por um produto interno indefinido gera uma teoria de contradomínios numéricos interessante e objecto de estudo recente [95, 96, 97]. Ao pretendermos mostrar o papel dos métodos algébricos na caracterização do contradomínio numérico clássico de alguns operadores lineares, com incidência em Física Quântica, fomos naturalmente conduzidos ao estudo do contradomínio numérico no âmbito de espaços com produto interno indefinido e aos resultados originais dos primeiros dois capítulos. Na parte final desta Dissertação, investigamos, ainda, algumas desigualdades espectrais de matrizes, envolvendo a entropia relativa quântica.

O primeiro capítulo apresenta uma visão global da teoria de contradomínios numéricos e de algumas das suas generalizações. Divide-se em quatro secções, cujos conteúdos passamos a pormenorizar. A primeira secção, dedicada ao contradomínio numérico clássico de operadores, inclui algumas propriedades bem conhecidas deste conjunto, entre as quais o Teorema do Contradomínio Elíptico e o Teorema de Toeplitz-Hausdorff. Esta compilação de resultados chave é não só útil, como pretende motivar a teoria que seguidamente se desenvolve.

Na segunda secção, introduz-se o conceito de contradomínio numérico em espaços de Krein, isto é, em espaços munidos de um produto interno indefinido. Restringe-se o estudo a espaços de dimensão finita n, no âmbito dos quais este conjunto se encontra naturalmente associado a uma matriz Hermítica não-singular H; razão pela qual se designa por *contradomínio numérico-H* de A e se denota por $W_H(A)$, sendo A uma matriz quadrada complexa de ordem n. Coligem-se algumas propriedades básicas de $W_H(A)$, que generalizam as antes apresentadas para o contradomínio numérico clássico, ou que realçam as diferenças que eventualmente ocorram. Por exemplo, $W_H(A)$ é pseudo-convexo, não necessariamente limitado nem fechado, por oposição a $W(A)$ que, para além de convexo, é compacto, em espaços de dimensão finita. Prossegue a exposição com um importante resultado relativo à curva geradora de fronteira de $W_H(A)$, que generaliza o Teorema de Murnagham-Kippenhahn [86, 105]. Uma observação crucial, relacionando o contradomínio numérico-H e o contradomínio numérico-J de matrizes, onde J é a matriz de inércia de H, permite simplificar drasticamente o estudo proposto. Seguem-se resultados envolvendo matrizes Hermíticas-J e normais-J, alguns dos quais oferecem respostas parciais a questões colocadas por Chi-Kwong Li, Nam-Kiu Tsing e Frank Uhlig [95]. A dominar a Subsecção 1.2.4, demonstramos o Teorema do Contradomínio Hiperbólico que caracteriza completamente o contradomínio numérico-J de uma matriz quadrada complexa de ordem dois, quando J é a matriz diagonal de entradas principais a unidade e o seu simétrico. Como aplicação, revelam-se algumas curvas algébricas hiperbólicas geradoras da fronteira do contradomínio numérico-J de matrizes por blocos, de ordem superior a dois.

Uma das generalizações mais famosas do contradomínio numérico clássico, dada uma matriz complexa C, é o contradomínio numérico-C. Este conceito motiva a definição do *contradomínio tracial-*C, H de uma matriz A, denotado por $W_C^H(A)$, e apresentado na terceira secção do primeiro capítulo. Investigam-se algumas suas formas especiais, tais como a forma elíptica ou hiperbólica para matrizes de dimensão dois, a de conjunto singular e a de subconjunto de uma recta. Como consequência do Teorema de Tarski, prova-se, seguindo a linha de abordagem de Leal Duarte [48], que a fronteira de $W_C^H(A)$ é uma união finita de arcos algébricos. Posto isto, estudam-se certos pontos fronteiros especiais de $W_C^J(A)$, os pontos angulosos, generalizando resultados previamente obtidos por Natália Bebiano [11].

Para finalizar o capítulo um, como variação do contradomínio tracial, define-se o *contradomínio determinantal-*C, H de uma matriz A, denotado por $\triangle_C^H(A)$. Apresentam-se alguns resultados relativos às formas especiais e aos pontos angulosos de $\triangle_C^J(A)$, paralelos aos antes apresentados para $W_C^J(A)$. Em particular, quando H é a matriz identidade, $\triangle_C^H(A)$ reduz-se ao contradomínio determinantal-C de A, cuja história está intimamente ligada à famosa conjectura de Marcus-Oliveira [100, 113], relativa ao determinante da soma de duas matrizes normais com valores próprios prescritos.

No segundo capítulo, investiga-se o contradomínio numérico clássico de alguns operadores descritos em termos de operadores de criação e destruição que surgem na Física Quântica. A iniciar o capítulo, algumas noções preliminares sobre classes simétricas de tensores clarificam o papel do m-ésimo espaço (completamente) simétrico e do m-ésimo espaço de Grassmann sobre um espaço de Hilbert. Estes revelam-se os espaços de estados adequados para descrever sistemas com m bosões e m fermiões, respectivamente.

A terceira secção, manifestamente a principal do segundo capítulo, centra a atenção nos operadores de emparelhamento, definidos em termos de operadores de criação e destruição.

Aqui, estudam-se algumas propriedades espectrais dos operadores de emparelhamento, utilizando uma transformação de Bogoliubov adequada, e obtêm-se importantes teoremas que caracterizam o contradomínio numérico clássico destes operadores ilimitados, quando restritos a determinados subespaços da álgebra simétrica sobre $\mathbb{C}^2$. Começando pelos operadores de emparelhamento bosónicos auto-adjuntos, obtém-se a descrição do contradomínio numérico de determinadas matrizes Hermíticas tridiagonais infinitas. No caso geral, prova-se que o contradomínio numérico destes operadores de emparelhamento toma a forma hiperbólica (possivelmente degenerada), bem como as suas representações matriciais tridiagonais infinitas. Fruto da relação que se estabelece entre o contradomínio numérico clássico destes operadores e o contradomínio numérico em espaços de Krein de uma certa matriz de ordem dois, estes resultados derivam do Teorema do Contradomínio Hiperbólico que ocupara lugar de destaque no primeiro capítulo. No que diz respeito a operadores de emparelhamento fermiónicos, ou seja, definidos na álgebra de Grassmann, o seu contradomínio numérico resulta simplesmente da representação matricial destes operadores na base canónica.

A terminar o segundo capítulo, investiga-se uma nova classe de operadores definidos em termos de operadores de criação e destruição bosónicos, restritos ao m-ésimo espaço simétrico sobre $\mathbb{C}^2$. A primeira derivação de matrizes complexas de ordem dois é dada por tais operadores lineares, que podem ser representados por matrizes tridiagonais finitas. Mostra-se que estes admitem contradomínios numéricos elípticos homotéticos de razão m, e generaliza-se o resultado ao contradomínio numérico-c, para $c \in \mathbb{R}^n$. Prova-se ainda a invariância do contradomínio numérico-c, $c \in \mathbb{R}^n$, perante a troca de duas entradas simétricas fora da diagonal principal de uma matriz tridiagonal, generalizando resultados recentes de Ethan Brown e Ilya Spitkovksy [36].

No terceiro capítulo, investigam-se algumas desigualdades matriciais com incidências em Física. Uma breve resenha histórica introduz as noções de entropia, segundo Shannon e von Neumann, a entropia relativa de Umegaki e a variante para operadores de Fujii e Kamei. Recorda-se a desigualdade de Furuta [58], uma generalização notável da desigualdade de Löwner-Heinz, à custa da qual se obtém uma nova majoração logarítmica, usando técnicas de Ando e Hiai. Esta majoração é utilizada para fornecer uma prova alternativa para um resultado conhecido de Hiai e Petz [77], complementado por Ando e Hiai [2], que estabelece um limite superior para a entropia relativa de Umegaki. Em particular, obtêm-se algumas desigualdades traciais logarítmicas, que se reescrevem em linguagem de entropia relativa.

As secções seguintes são dedicadas à desigualdade de Peierls-Bogoliubov, que desempenha um papel importante na estimativa de um valor aproximado para a função de partição de um sistema físico, e à desigualdade termodinâmica. Generalizamos a primeira, fornecemos uma prova puramente algébrica para a segunda. Recorrendo a uma variante da desigualdade de Golden-Thompson, obtemos ainda uma versão generalizada da desigualdade termodinâmica, envolvendo a média da potência-α de matrizes.

Quase a finalizar, demonstramos uma cadeia de equivalências, contendo algumas das desigualdades antes obtidas. Em particular, estabelecemos uma relação entre a desigualdade termodinâmica, a desigualdade de Peierls-Bogoliubov e a positividade da entropia relativa para matrizes de densidade.

Acrescentamos, em jeito de considerações finais, alguns problemas em aberto que decorreram da nossa investigação, ilustrativos de potenciais linhas de rumo nesta empolgante área da teoria de matrizes e teoria de operadores lineares.

Preocupámo-nos em orientar a exposição de modo lógico e sequencial. Perante a vastidão e riqueza do tema, recorremos a uma bibliografia vasta. Para tornar o texto auto-contido, sintetizamos alguns conteúdos preliminares, no início de cada capítulo, e incluimos demonstrações de resultados auxiliares, mesmo que não da nossa autoria, mas que são reveladores das ferramentas e técnicas usadas. Perante diferentes provas conhecidas de um mesmo resultado, optamos pelas que consideramos mais elegantes e procuramos, quando possível, proporcionar versões mais simplificadas. A referência bibliográfica que acompanha a demonstração identifica a fonte em que nos baseamos. Para alguns teoremas conhecidos, apresentamos novas provas.

Os resultados originais desta Dissertação foram divulgados, na sua maioria, à comunidade científica internacional em conferências da área [19]; alguns dos quais estão já publicados [20, 21, 22] em revistas da especialidade, outros submetidos e aceites, encontram-se em fase de publicação [23, 25].

Neste ponto, importa recordar os diferentes estádios por que passou esta Dissertação, nomeadamente, em relação aos conjuntos $W_H(A)$, $W_C^H(A)$ e $\triangle_C^H(A)$, que sofreram alteração na definição inicial [22], onde eram denotados por $V_H(A)$, $V_{H,C}(A)$ e $D_{H,C}(A)$, respectivamente. Esta nova forma parece adaptar-se melhor ao propósito de evidenciar as similitudes e diferenças com os conjuntos $W(A)$, $W_C(A)$ e $\triangle_C(A)$ que generalizam. Também reflexo desta evolução é o enunciado do Teorema do Contradomínio Hiperbólico (e suas versões para o caso tracial e determinantal) e alguns resultados demonstrados nas Subsecções 1.3.1 e 1.4.1, em condições um pouco mais gerais do que em [22].

Apraz-nos saber que parte do trabalho por nós desenvolvido, no âmbito das desigualdades de matrizes, foi muito recentemente alvo de citação por parte de outros investigadores [57, 64, 65, 133]. Esperamos que este trabalho seja um pequeno contributo para a área dos contradomínios numéricos, das desigualdades matriciais e suas potenciais aplicações em Física.

Capítulo 1

Contradomínios Numéricos

A matemática pura é, à sua maneira, a poesia das ideias lógicas.

Albert Einstein

1.1 Contradomínio Numérico Clássico

Seja $\mathcal{H}$ um espaço de Hilbert complexo munido de um produto interno $\langle \cdot, \cdot \rangle$ e seja A um operador linear definido em $\mathcal{H}$. O *contradomínio numérico do operador* A é o subconjunto do plano complexo de Argand que se denota e define por

$$W(A) \;=\; \{\langle Ax, x\rangle :\; x \in \mathcal{H},\; \langle x, x\rangle = 1\}. \tag{1.1}$$

Otto Toeplitz [127] e Felix Hausdorff [75] referiram-se a este conjunto como o *Wertvorrat* de uma forma bilinear, a comunidade russa chamou-lhe domínio de Hausdorff, outros autores optaram pela denominação *field of values* (campo de valores) e Marshall Stone [123], na sua obra influente em teoria de operadores, sugeriu a terminologia *numerical range* (contradomínio numérico). Nesta competição de designações, a letra inicial da primeira ficou consagrada na notação $W(A)$, tendo apenas as duas últimas sobrevivido na nomenclatura corrente.

Supõe-se, ao longo desta Dissertação, que o espaço de Hilbert é separável, ou equivalentemente, que tem dimensão numerável [45, 4.16].

Se $\mathcal{H}$ é de dimensão finita n, escolhida uma base ortonormada no espaço $\mathcal{H}$, podem obviamente representar-se os vectores de $\mathcal{H}$ por n-uplos complexos e o operador A por um elemento da álgebra M_n das matrizes quadradas complexas de ordem n. Considerando $\mathbb{C}^n$ munido do produto interno euclidiano

$$\langle x, y\rangle \;=\; y^* x, \qquad x,\, y \,\in\, \mathbb{C}^n,$$

define-se o *contradomínio numérico da matriz* $A \in M_n$ por

$$W(A) \;=\; \{x^* A\, x:\; x \in \mathbb{C}^n,\; x^* x = 1\}. \tag{1.2}$$

Neste contexto, alguns autores [81] preferem manter a designação *campo de valores* de A e usam, em alternativa, a notação $F(A)$[1], contudo na literatura actual é irrefutável a supremacia do termo *contradomínio numérico* de A e da notação $W(A)$.

Se $\mathcal{H}$ é de dimensão infinita, verificam-se diferenças entre algumas propriedades do conjunto (1.1) e da sua versão matricial (1.2). Estas diferenças podem mesmo ser assinaláveis, quando se trata de operadores lineares ilimitados. Aprofundaremos este assunto no segundo capítulo.

Até final desta secção, consideram-se A e B operadores lineares limitados definidos no espaço de Hilbert $\mathcal{H}$, e denotam-se por I e $I_{\mathcal{V}}$ os operadores identidade em $\mathcal{H}$ e num seu subespaço $\mathcal{V}$, respectivamente.

Realçam-se algumas propriedades básicas do contradomínio numérico clássico, nomeadamente, a inclusão espectral:

W1. $\sigma(A) \subseteq \overline{W(A)}$[2], onde $\sigma(A)$ denota o espectro de A;

a propriedade da translação e multiplicação por um escalar:

W2. $W(\alpha I + \beta A) = \alpha + \beta\, W(A)$, quaisquer que sejam os escalares $\alpha, \beta \in \mathbb{C}$;

a invariância unitária:

W3. $W(U^* A U) = W(A)$, qualquer que seja o operador unitário U;

e, mais geralmente, dado $1 \le r \le n$ e $\mathcal{V}$ um subespaço vectorial de $\mathcal{H}$ de dimensão r, a propriedade da projecção isométrica:

W3'. $W(P^* A P) \subseteq W(A)$, para qualquer operador $P\colon \mathcal{V} \to \mathcal{H}$ tal que $P^* P = I_{\mathcal{V}}$, com igualdade se $r = n$;

a acção do operador adjunto de A:

W4. $W(A^*) = \{\bar{z} : z \in W(A)\}$;

e a subaditividade:

W5. $W(A + B) \subseteq W(A) + W(B)$.

A inclusão espectral foi originalmente provada, em 1929, por Aurel Wintner [132]. (Para outra demonstração ver, por exemplo, Paul R. Halmos [73, Solução 214].) As

[1] A letra F deriva da terminologia inglesa *field of values*.

[2] $\overline{K}$ denota o fecho topológico do conjunto K.

propriedades W2-W5 são consequência imediata da definição. (Para mais pormenores, consultar [71] e [81, capítulo 1].)

Define-se o *raio numérico* de A por $r(A) = \sup\{|z| : z \in W(A)\}$. Verifica-se que $r(A)$ é uma norma equivalente à *norma do operador*

$$\|A\| \; = \; \sup_{\|x\|=1} \|Ax\|,$$

pois $r(A) \leq \|A\| \leq 2\, r(A)$ [71, p.9], donde se conclui que se $r(A) = 0$ então $A = 0$.

Este último facto, em conjunto com W2, permite caracterizar os operadores cujo contradomínio numérico é um conjunto singular:

W6. $W(A) = \{\lambda\}$ se e só se $A = \lambda I$, para determinado $\lambda \in \mathbb{C}$;

e, em conjunto com W4 e W5, os operadores cujo contradomínio numérico é um subconjunto do eixo real:

W7. $W(A) \subseteq \mathbb{R}$ se e só se A é um operador auto-adjunto, isto é, $A = A^*$.

Atendendo à *decomposição cartesiana* de A, dada por $A = \mathrm{Re}A + i\,\mathrm{Im}A$, em que

$$\mathrm{Re}A \; = \; \frac{A + A^*}{2} \qquad \text{e} \qquad \mathrm{Im}A \; = \; \frac{A - A^*}{2i}$$

são operadores auto-adjuntos, e às propriedades W5 e W2, tem-se a seguinte inclusão:

W8. $W(A) \subseteq W(\mathrm{Re}A) + i\, W(\mathrm{Im}A)$.

Além disso, valem as relações:

W9. $\mathrm{Re}\, W(A) = W(\mathrm{Re}A)$ e $\mathrm{Im}\, W(A) = W(\mathrm{Im}A)$,

em que $\mathrm{Re}\, K$ e $\mathrm{Im}\, K$ denotam as projecções ortogonais do conjunto K nos eixos real e imaginário, respectivamente.

Em virtude de $W(A)$ ser a imagem da superfície da esfera unitária pela aplicação contínua $x \mapsto \langle Ax, x\rangle$, tem-se que

W10. $W(A)$ é um conjunto compacto e conexo, se o espaço $\mathcal{H}$ é de dimensão finita.

Se $A \in M_n$, $1 \leq r \leq n$ e $P \in M_{n,r}$ se obtém da matriz identidade I_n por eliminação de $n - r$ colunas, então P^*AP é uma submatriz principal de A, obtendo-se de W3' a propriedade da inclusão para submatrizes:

W11. $W(B) \subseteq W(A)$, qualquer que seja a submatriz principal B da matriz A.

1.1.1 Teorema do Contradomínio Elíptico

A primeira demonstração conhecida do *Teorema do Contradomínio Elíptico* data de 1932 e é devida a Francis D. Murnaghan [105]. Mais tarde, William F. Donoghue [47] descreve a fronteira elíptica do contradomínio numérico de um operador linear A definido num espaço de Hilbert bidimensional, como a envolvente a uma família de circunferências. Desde então, surgiram na literatura várias outras provas (vide [71, 81, 118, 122] e suas referências); destacamos a mais recente de Chi-Kwong Li [93] que, de uma forma elegante, reduz o estudo ao caso circular.

Lema 1.1.1 *O contradomínio numérico da matriz*

$$A_1 \;=\; \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}. \tag{1.3}$$

é um disco circular centrado na origem e de diâmetro um.

Demonstração: Pela definição de contradomínio numérico, tem-se

$$W(A_1) \;=\; \left\{ \bar{x}_1 x_2 : \; x_1, x_2 \in \mathbb{C}, \; |x_1|^2 + |x_2|^2 = 1 \right\}. \tag{1.4}$$

Tomando $r = |x_1|^2$ e $\theta = \arg x_2 - \arg x_1$, pode reescrever-se (1.4) na forma

$$W(A_1) \;=\; \bigcup_{r\in[0,1]} \left\{ \sqrt{r(1-r)}\,\mathrm{e}^{i\theta} : \; \theta \in [0, 2\pi) \right\}.$$

Quando θ varia de 0 a 2π, os elementos de $W(A_1)$, para r fixo, descrevem uma circunferência centrada na origem e de raio $\sqrt{r(1-r)}$. À medida que r varia de 0 a 1, os raios das circunferências percorrem todos os valores de 0 a $1/2$ e $W(A_1)$, sendo a união desta família de curvas, é um disco circular. ■

Teorema 1.1.1 (Teorema do Contradomínio Elíptico, 1932) *Seja $A \in M_2$ de valores próprios α_1 e α_2. O contradomínio numérico de A é um disco elíptico (possivelmente degenerado), de focos α_1 e α_2, cujos eixos maior e menor têm comprimento*

$$\sqrt{\mathrm{Tr}\,(A^*A) - 2\mathrm{Re}\,(\overline{\alpha}_1\alpha_2)} \qquad e \qquad \sqrt{\mathrm{Tr}\,(A^*A) - |\alpha_1|^2 - |\alpha_2|^2}, \tag{1.5}$$

respectivamente.

Demonstração [93]**:** Pelo Teorema da Triangularização de Schur [80, p.79], toda a matriz $A \in M_2$ é unitariamente semelhante a uma matriz triangular superior do tipo

$$B \;=\; \begin{bmatrix} \alpha_1 & b \\ 0 & \alpha_2 \end{bmatrix}, \qquad b \in \mathbb{C},$$

e, pela invariância unitária W3, vale $W(A) = W(B)$.

I. Se A é normal, então $b = 0$ e $W(B) = \{\alpha_1 r + \alpha_2(1-r) : 0 \le r \le 1\}$ é o segmento de recta fechado de extremos α_1 e α_2 (que se reduz ao ponto α_1, quando A é escalar).

II. Seja A não normal, isto é, $b \neq 0$, e $\alpha_1 = \alpha_2$. Então $B = \alpha_1 I_2 + bA_1$, onde A_1 é a matriz (1.3). A partir do Lema 1.1.1 e recorrendo à propriedade W2, conclui-se que $W(B) = \alpha_1 + b\,W(A_1)$ é um disco circular centrado em α_1 e de diâmetro $|b|$.

III. Se A é não normal e $\alpha_1 \neq \alpha_2$, então $B = \frac{1}{2}\,\mathrm{Tr}\,(A)\,I_2 + \frac{1}{2}\,(\alpha_1 - \alpha_2)\,V^*CV$, onde

$$C \;=\; \begin{bmatrix} 1 & 2|c| \\ 0 & -1 \end{bmatrix} \qquad \text{e} \qquad V \;=\; \begin{bmatrix} 1 & 0 \\ 0 & \mathrm{e}^{i\arg c} \end{bmatrix}, \qquad c \;=\; \frac{b}{\alpha_1 - \alpha_2} \;\neq\; 0.$$

Sendo V uma matriz unitária, as propriedades W2 e W3 garantem que

$$W(B) \;=\; \frac{1}{2}\,\mathrm{Tr}\,(A) + \frac{1}{2}\,(\alpha_1 - \alpha_2)\,W(C), \tag{1.6}$$

pelo que basta estudar o contradomínio numérico de C. Verifica-se que a matriz

$$D \;=\; |c|\,\mathrm{Re}\,C + i\sqrt{1+|c|^2}\,\mathrm{Im}\,C$$

tem ambos os valores próprios nulos e, pelo Teorema da Triangularização de Schur, existe $U \in M_2$ unitária tal que $U^*DU \;=\; 2|c|\sqrt{1+|c|^2}A_1$, onde A_1 é a matriz (1.3). Pelo Lema 1.1.1, conclui-se que $W(D)$ é um disco circular centrado na origem e de raio $|c|\sqrt{1+|c|^2}$. Claramente,

$$|c|\,x + i\sqrt{1+|c|^2}\,y \,\in\, W(D) \qquad \Leftrightarrow \qquad x + iy \,\in\, W(C).$$

Conhecida a forma circular e o raio de $W(D)$, tem-se $x = \sqrt{1+|c|^2}\cos\theta$ e $y = |c|\sin\theta$, $\theta \in [0, 2\pi)$. Portanto, $W(C)$ é um disco elíptico de focos 1 e -1, de eixos maior e menor de comprimento $2\sqrt{1+|c|^2}$ e $2|c|$, respectivamente. Além disso, sendo $W(B)$ uma translação do conjunto $W(C)$ previamente multiplicado por $b/(2c)$, conforme (1.6), é claro que $W(B)$ é um disco elíptico; os seus focos são os valores próprios α_1 e α_2, e os eixos maior e menor têm comprimento $\sqrt{|\alpha_1 - \alpha_2|^2 + |b|^2}$ e $|b|$, respectivamente. Para terminar, observa-se que $\mathrm{Tr}\,(A^*A) = |\alpha_1|^2 + |\alpha_2|^2 + |b|^2$, o que permite reescrever os comprimentos dos eixos da elipse na forma (1.5). ■

A *norma euclideana* ou *de Frobenius* de $A \in M_n$ é dada por $\|A\|_2 = \sqrt{\mathrm{Tr}(A^*A)}$. O contradomínio numérico de $A \in M_2$ é, assim, um disco elíptico completamente determinado por $\|A\|_2$ e pelos valores próprios de A.

Em especial, se A é um operador linear definido num espaço de Hilbert bidimensional, cujos valores próprios α_1 e α_2 são distintos, então W. F. Donoghue [47] provou que a excentricidade da elipse que delimita $W(A)$ é dada por $\sin\varphi$, onde

$$\varphi \;=\; \arccos\frac{|\langle x_1, x_2\rangle|}{\|x_1\|\|x_2\|}, \qquad x_i \in E_{\alpha_i}, \;\; i = 1, 2,$$

é o ângulo reduzido entre os dois subespaços próprios E_{α_1} e E_{α_2} de A, associados a α_1 e α_2, respectivamente.

Apesar da descrição geométrica do contradomínio numérico de $A \in M_n$ ser particularmente simples para $n = 2$, em geral, pode revelar-se complicada. No caso $n = 3$, investigado por D. Keeler, L. Rodman e I. Spitkovsky [85], o contradomínio numérico pode tomar a forma elíptica, cónica ou triangular, ovular ou com uma porção plana na fronteira. Se $n \geq 4$, conhece-se, apenas, a forma geométrica do contradomínio numérico de algumas classes especiais de matrizes e desconhecem-se quais os conjuntos convexos e compactos do plano que podem ocorrer como o contradomínio numérico de matrizes.

1.1.2 Teorema de Toeplitz-Hausdorff

Em 1918, Otto Toeplitz [127] mostrou que a fronteira do contradomínio numérico é uma curva convexa, mas não excluiu a possibilidade de buracos ocorrerem no seu interior. Seis meses depois, Felix Hausdorff [75] provou que o conjunto é simplesmente conexo. Desde então, vários autores (consultar, por exemplo, [46, 68, 81]) apresentaram demonstrações alternativas para a convexidade de $W(A)$. Neste percurso, o Teorema do Contradomínio Elíptico desempenha um papel central, já que a maioria das provas reduz o estudo da convexidade de $W(A)$ ao caso bidimensional. Tal redução é viável, mesmo quando o espaço subjacente é de dimensão infinita.

Seja $\mathcal{V}$ um subespaço vectorial de $\mathcal{H}$ e $P_\mathcal{V}$ a projecção ortogonal de $\mathcal{H}$ em $\mathcal{V}$. Dado um operador linear A em $\mathcal{H}$, o operador $P_\mathcal{V}A$ restrito ao subespaço $\mathcal{V}$ diz-se a *compressão* do operador A a $\mathcal{V}$.

Lema 1.1.2 *Se A é um operador linear em $\mathcal{H}$, então $W(A)$ contém o contradomínio numérico de todas as compressões de A.*

Demonstração [122]**:** Seja $\mathcal{V}$ um subespaço vectorial de $\mathcal{H}$ e designe-se por $A_\mathcal{V}$ a compressão de A a $\mathcal{V}$. Dado $x \in \mathcal{V}$, tem-se

$$\langle A_\mathcal{V}\, x, x\rangle \;=\; \langle P_\mathcal{V} A x, x\rangle \;=\; \langle Ax, P_\mathcal{V}\, x\rangle \;=\; \langle Ax, x\rangle \;\in\; W(A),$$

logo $W(A_\mathcal{V}) \subseteq W(A)$. ■

Teorema 1.1.2 (Teorema de Toeplitz-Hausdorff, 1918) *Se A é um operador linear em $\mathcal{H}$, então $W(A)$ é um conjunto convexo.*

Demonstração [47]**:** Sejam z, w dois pontos distintos de $W(A)$ e $\mathcal{L}$ o segmento de recta que os une. Existem vectores unitários $x, y \in \mathcal{H}$ tais que $z = \langle Ax, x\rangle$ e $w = \langle Ay, y\rangle$. Seja A_{xy} a compressão de A ao subespaço vectorial gerado por x e y (de dimensão dois, caso contrário $z = w$). Pelo Teorema do Contradomínio Elíptico, $W(A_{xy})$ é um disco elíptico, que por conter os pontos z e w, também contém o segmento de recta $\mathcal{L}$. Atendendo ao Lema 1.1.2, tem-se

$$W(A) \;\supseteq\; W(A_{xy}) \;\supseteq\; \mathcal{L},$$

o que prova que $W(A)$ é convexo. ∎

O Teorema de Toeplitz-Hausdorff continua, ainda hoje, a ser uma fonte de inspiração para muitos investigadores (refira-se, por exemplo, uma prova muito recente deste teorema [72] obtida, num âmbito mais geral, usando métodos de geometria e topologia diferencial).

Estabelecida a convexidade, distinguem-se mais algumas propriedades do contradomínio numérico, como a da soma directa [81, p.12]:

W12. $W(A \oplus B) = \mathrm{co}\,(W(A) \cup W(B))$, quaisquer que sejam $A, B \in M_n$,

em que $\mathrm{co}\,K$ denota o invólucro convexo do conjunto K. Além disso, a propriedade da inclusão espectral admite o seguinte refinamento:

W1'. $\mathrm{co}\,\sigma(A) \subseteq \overline{W(A)}$, qualquer que seja o operador linear A em $\mathcal{H}$,

ocorrendo igualdade de conjuntos em W1', se A for um operador normal [71, 123]. Em particular, verifica-se que:

W13. $W(A) = [\lambda_n, \lambda_1]$, se $A \in M_n$ é Hermítica de valores próprios $\lambda_1 \geq \cdots \geq \lambda_n$.

Dada $A \in M_n$, atendendo às propriedades W8, W13 e ao facto de $\mathrm{Re}A$ e $\mathrm{Im}A$ serem matrizes Hermíticas, conclui-se que $W(A)$ está contido no rectângulo com vértices os valores próprios máximo e mínimo de $\mathrm{Re}A$ e $\mathrm{Im}A$.

Designa-se *recta de suporte* de um subconjunto convexo K do plano complexo toda a recta que intersecta K em pelo menos um ponto e que define dois semi-planos, um dos quais não contém ponto algum de K. Denota-se a fronteira de K por ∂K.

Claramente, se $\lambda_1(\mathrm{Re}A)$ e $\lambda_n(\mathrm{Re}A)$ denotam o valor próprio máximo e mínimo de $\mathrm{Re}A$, respectivamente, então

$$\mathcal{L}_1 \;=\; \{\lambda_n(\mathrm{Re}A) + iy : \; y \in \mathbb{R}\} \qquad \text{e} \qquad \mathcal{L}_2 \;=\; \{\lambda_1(\mathrm{Re}A) + iy : \; y \in \mathbb{R}\}$$

são as rectas de suporte vertical esquerda e direita de $W(A)$. Se x é um vector próprio unitário de ReA associado ao seu valor próprio máximo (mínimo) e se $z = \langle Ax, x\rangle$, então $z \in \mathcal{L}_2$ ($z \in \mathcal{L}_1$); portanto z é um ponto da fronteira de $W(A)$.

Proposição 1.1.1 *Seja A um operador linear em $\mathcal{H}$ e $\theta \in [0, 2\pi)$. Supondo que existe valor próprio máximo (mínimo) de* $\mathrm{Re}\,(\mathrm{e}^{i\theta}A)$ *e que x_θ é o correspondente vector próprio unitário, então $\langle Ax_\theta, x_\theta\rangle \in \partial W(A)$. Em particular, se $A \in M_n$, então*

$$W(A) \;=\; \mathrm{co}\,\{\langle Ax_\theta, x_\theta\rangle :\; 0 \leq \theta < 2\pi\}.$$

Demonstração: Seja

$$\Omega \;=\; \{\theta \in [0, 2\pi) :\; \mathrm{Re}\,(\mathrm{e}^{i\theta}A) \text{ admite valor próprio máximo}\}$$

e $A_\theta = \mathrm{e}^{i\theta}A$, $\theta \in \Omega$. Tendo em conta as hipóteses e a observação que precedeu este resultado, é claro que $\langle A_\theta\, x_\theta, x_\theta\rangle \in \partial W(A_\theta)$. Pela propriedade W3 relativa à multiplicação escalar, tem-se $\mathrm{e}^{-i\theta}W(A_\theta) = W(A)$, donde $\langle Ax_\theta, x_\theta\rangle = \mathrm{e}^{-i\theta}\langle A_\theta\, x_\theta, x_\theta\rangle \in \partial W(A)$. Quando $A \in M_n$, obviamente $\Omega = [0, 2\pi)$ e a conclusão decorre da convexidade de $W(A)$. ■

Um *ponto anguloso* de um subconjunto convexo K do plano complexo é um ponto na fronteira de K que é vértice de um sector, contendo K, de medida angular inferior a π radianos. É claro que por um ponto anguloso pertencente a K passam pelo menos duas rectas de suporte de K.

A fronteira de um conjunto convexo é uma curva diferenciável, excepto quando muito num conjunto numerável de pontos angulosos [122]. Em virtude da convexidade do contradomínio numérico, todos os pontos não diferenciáveis de $\partial W(A)$ são pontos angulosos. Em Física, os pontos angulosos correspondem a estados particularmente importantes. Por exemplo, o estado de vácuo produz um ponto anguloso do contradomínio numérico do operador energia [18].

Além disso, existe uma ligação entre os pontos angulosos do contradomínio numérico de um operador e o seu espectro. Este resultado, originalmente formulado por Rudolf Kippenhahn [86], é frequentemente atribuído a William F. Donoghue [47] que o restabeleceu em condições mais gerais.

Teorema 1.1.3 *Seja A um operador linear em $\mathcal{H}$. Se $z \in W(A)$ é um ponto anguloso de $W(A)$, então z é um valor próprio de A.*

Demonstração [47]**:** Sejam $x \in \mathcal{H}$ um vector unitário tal que $z = \langle Ax, x\rangle$ e $y \in \mathcal{H} \setminus \{0\}$. Seja A_{xy} a compressão de A ao subespaço vectorial gerado por x e y. Atendendo ao

Teorema do Contradomínio Elíptico e ao Lema 1.1.2, a fronteira de $W(A_{xy})$ é uma elipse (possivelmente degenerada) contida em $W(A)$. Como z é um ponto anguloso de $W(A)$ e $z \in W(A_{xy})$, então z é necessariamente um ponto anguloso de $W(A_{xy})$. Assim sendo, o disco elíptico $W(A_{xy})$ degenera num segmento de recta, de extremo z, ou no ponto z. Em ambos os casos, z é um valor próprio da compressão A_{xy} e, consequentemente, um valor próprio do operador A, de vector próprio associado x. ■

1.2 Contradomínio Numérico em Espaços de Krein

Em álgebra linear, os conceitos de comprimento, ângulo e ortogonalidade definem-se tradicionalmente à custa de um produto interno (definido). Um produto interno *indefinido* satisfaz todas as condições de um produto interno (definido), excepto possivelmente a positividade. A utilização de um produto interno indefinido pode produzir alterações substanciais na geometria, referindo-se, entre as várias aplicações, a sua importância na teoria da relatividade. Analisar o efeito de tal substituição no contradomínio numérico de matrizes é o objectivo desta segunda secção.

Seja $\mathcal{K}$ um espaço de Hilbert complexo dotado de um produto interno $\langle \cdot, \cdot \rangle$ e seja H um operador linear em $\mathcal{K}$, limitado, auto-adjunto e invertível. A fórmula

$$[x, y] \;=\; \langle Hx, y \rangle, \qquad x,\, y \,\in\, \mathcal{K}, \tag{1.7}$$

determina um *produto interno indefinido* em $\mathcal{K}$. Reciprocamente, todo o produto interno indefinido $[\cdot, \cdot]$ em $\mathcal{K}$ é induzido por um único operador linear H, definido em $\mathcal{K}$, limitado, auto-adjunto e invertível, tal que (1.7) se verifica [66]. Por este motivo, escreve-se $\langle x, y \rangle_H$ em vez de $[x, y]$. O espaço de Hilbert complexo $\mathcal{K}$ munido de um produto interno indefinido diz-se um *espaço de Krein*.

Sendo A um operador linear limitado num espaço de Krein $\mathcal{K}$, a noção de contradomínio numérico de A no espaço de Krein $\mathcal{K}$ está naturalmente associada a um operador linear limitado, auto-adjunto e invertível H, definido nesse mesmo espaço $\mathcal{K}$. Atribui-se justamente a designação de *contradomínio numérico-H* de A a este subconjunto do plano complexo de Argand que se denota e define por

$$W_H(A) \;=\; \left\{ \frac{\langle Ax, x \rangle_H}{\langle x, x \rangle_H} :\; x \in \mathcal{K},\; \langle x, x \rangle_H \neq 0 \right\}.$$

Se H é o operador identidade, então $W_H(A)$ reduz-se ao contradomínio numérico clássico $W(A)$. O conjunto generalizado $W_H(A)$ marca uma nova era de investigação, impulsionada por Chi-Kwong Li, Nam-Kiu Tsing e Frank Uhlig [95] num artigo que, em 1996, inaugura a revista *Electronic Journal of Linear Algebra*.

Restringimos o nosso estudo a espaços de Krein $\mathcal{K}$ de dimensão finita n. Assim, ao longo desta segunda secção, identifica-se $\mathcal{K}$ com $\mathbb{C}^n$, munido do produto interno indefinido

$$\langle x, y\rangle_H \;=\; y^* H\, x, \qquad x,\, y\, \in\, \mathbb{C}^n,$$

onde $H \in M_n$ é uma matriz Hermítica não-singular, e dirigimos a atenção para o contradomínio numérico-H de matrizes $A \in M_n$.

Por conveniência, consideram-se os conjuntos relacionados

$$W_H^{\pm}(A) \;=\; \{\langle Ax, x\rangle_H :\; x \in \mathbb{C}^n,\; \langle x, x\rangle_H = \pm 1\}$$

que designaremos por *contradomínio numérico-H positivo* ou *negativo*[3] de $A \in M_n$. Se $H = \pm I_n$, então $\pm W_H^{\pm}(A) = W(A)$ e $W_H^{\mp}(A) = \emptyset$. Facilmente se verifica que

$$W_{-H}^{+}(A) \;=\; -W_H^{-}(A) \qquad \text{e} \qquad W_H(A) \;=\; W_H^{+}(A) \cup W_{-H}^{+}(A).$$

Alguns autores [96, 97] preferiram adoptar o conjunto $W_H^+(A)$ ao generalizar o contradomínio numérico a espaços de Krein, por $W_H^+(A)$ decorrer naturalmente de (1.2) por troca do produto escalar euclidiano por um indefinido. Contudo a falta de simetria deste último conjunto pode limitar a sua utilidade.

1.2.1 Propriedades Básicas

Doravante, H_n denota o espaço vectorial real das matrizes Hermíticas de M_n e $H \in H_n$ é não-singular. Introduzimos, antes de mais, a noção de matriz adjunta relativamente ao produto interno indefinido induzido por H. A *adjunta-H* de $A \in M_n$ é a matriz $A^{[*]}$ unicamente determinada pela relação

$$\langle Ax, y\rangle_H \;=\; \langle x, A^{[*]}y\rangle_H, \qquad x,\, y\, \in\, \mathbb{C}^n,$$

que se pode exprimir explicitamente em termos das matrizes A e H por

$$A^{[*]} \;=\; H^{-1}A^*H.$$

As definições seguintes surgem, assim, naturalmente. Uma matriz $A \in M_n$ diz-se *Hermítica-H* se $A = A^{[*]}$. Uma matriz $N \in M_n$ diz-se *normal-H* se $NN^{[*]} = N^{[*]}N$. Uma matriz $U \in M_n$ diz-se *unitária-H* se U é invertível e $U^{-1} = U^{[*]}$. Dada $H \in H_n$ uma matriz não-singular prescrita, o conjunto das matrizes unitárias-H forma um grupo.

[3] A notação $W_H^{\pm}(A)$ deve ser entendida da seguinte forma: quando x percorre o conjunto dos vectores para os quais $\langle x, x\rangle_H = 1$, define-se o contradomínio numérico-H positivo $W_H^+(A)$; quando x percorre os vectores para os quais $\langle x, x\rangle_H = -1$, define-se o contradomínio numérico-H negativo $W_H^-(A)$.

Um subespaço E de $\mathbb{C}^n$ diz-se *isotrópico-H* se $\langle x, x\rangle_H = 0$, para todo o $x \in E$.

Observa-se que o espectro de uma matriz U unitária-H é simétrico relativamente ao círculo unitário $\mathcal{D}$, isto é, se $\lambda \in \sigma(U)$ então $\bar{\lambda}^{-1} \in \sigma(U)$ [66, Proposição 2.7] e E_λ, subespaço próprio de U associados a $\lambda \notin \mathcal{D}$, é isotrópico-$H$ [66, p.29].

Por oposição à noção de vector isotrópico-H, um vector $x \in \mathbb{C}^n$ tal que $\langle x, x\rangle_H \neq 0$ denomina-se *anisotrópico-H*. Denotando por $\sigma_H(A)$ ($\sigma_H^{\pm}(A)$) o conjunto dos valores próprios de A que têm vectores próprios anisotrópicos-H (vectores próprios $x \in \mathbb{C}^n$ tais que $\langle x, x\rangle_H = \pm 1$), verifica-se que

$\mathbf{W_H1.}$ $\sigma_H(A) \subseteq W_H(A)$; $\mathbf{W_H1'.}$ $\sigma_H^{\pm}(A) \subseteq \pm W_H^{\pm}(A)$.

Se $H = I_n$, então $\sigma_H(A) = \sigma(A)$ e $\mathrm{W_H}1$ reduz-se à inclusão espectral do contradomínio numérico clássico.

Listamos, agora, algumas propriedades do contradomínio numérico-H que são consequência directa da definição, começando pela referente à translação e multiplicação escalar:

$\mathbf{W_H2.}$ $W_H(\alpha I_n + \beta A) = \alpha + \beta\, W_H(A)$, para quaisquer escalares $\alpha, \beta \in \mathbb{C}$;

passando à invariância unitária-H:

$\mathbf{W_H3.}$ $W_H(U^{[*]}AU) = W_H(A)$, qualquer que seja a matriz U unitária-H;

à acção da matriz adjunta-H de A:

$\mathbf{W_H4.}$ $W_H(A^{[*]}) = \{\bar{z} : z \in W_H(A)\}$;

e à propriedade subaditiva:

$\mathbf{W_H5.}$ $W_H(A + B) \subseteq W_H(A) + W_H(B)$.

Importa realçar que as mesmas propriedades são válidas ao substituir W_H por $W_H^{\pm}$, excepto $\mathrm{W_H}2$. Em alternativa, verifica-se

$\mathbf{W_H2'.}$ $W_H^{\pm}(\alpha I + \beta A) = \pm\alpha + \beta W_H^{\pm}(A)$, para quaisquer escalares $\alpha, \beta \in \mathbb{C}$.

A caracterização das matrizes cujo contradomínio numérico-H é um conjunto singular [95, Teorema 2.3 b] revela um claro paralelismo com o caso clássico:

$\mathbf{W_H6.}$ $W_H(A) = \{\lambda\}$ se e só se $A = \lambda I_n$, para determinado $\lambda \in \mathbb{C}$;

bem como a caracterização das matrizes cujo contradomínio numérico-H é um subconjunto do eixo real [95, Corolário 2.5]:

$\mathbf{W_H7.}$ $W_H(A) \subseteq \mathbb{R}$ se e só se A é um matriz Hermítica-H.

Estas duas últimas caracterizações valem ainda para os conjuntos $W_H^{\pm}(A)$, desde que satisfeitas as respectivas condições $\sigma(H) \cap \mathbb{R}^{\pm} \neq \emptyset$.

Qualquer matriz $A \in M_n$ pode decompor-se na forma $A = \mathrm{Re}^H A + i\, \mathrm{Im}^H A$, em que

$$\mathrm{Re}^H A \;=\; \frac{A + A^{[*]}}{2} \qquad \text{e} \qquad \mathrm{Im}^H A \;=\; \frac{A - A^{[*]}}{2i}$$

são matrizes Hermíticas-H. Quando $H = I_n$, esta reduz-se à decomposição cartesiana de A. Atendendo às propriedades $\mathrm{W_H}5$ e $\mathrm{W_H}2$, tem-se

$\mathbf{W_H}$8. $W_H(A) \subseteq W_H(\mathrm{Re}^H A) + i\, W_H(\mathrm{Im}^H A)$;

e, claramente, valem as igualdades:

$\mathbf{W_H}$9. $\mathrm{Re}\, W_H(A) = W_H(\mathrm{Re}^H A)$ e $\mathrm{Im}\, W_H(A) = W_H(\mathrm{Im}^H A)$.

O *contradomínio numérico conjunto* de três[4] matrizes $F, G, H \in H_n$ é o subconjunto de $\mathbb{R}^3$ definido por

$$W(F, G, H) \;=\; \{(x^* F x, x^* G x, x^* H x) :\ x \in \mathbb{C}^n,\ x^* x = 1\}, \tag{1.8}$$

Au-Yeung e Tsing [6] mostraram que $W(F, G, H)$ é um conjunto convexo, se $n > 2$ (vide [72] para outra abordagem), o qual se reduz à superfície de um elipsóide (eventualmente degenerado), se $n = 2$.

Um subconjunto S de $\mathbb{R}^3$ diz-se um *cone* se, para todo o $\alpha > 0$ e qualquer que seja o $x \in S$, se verifica $\alpha\, x \in S$. Chi-Kwong Li, Nam-Kiu Tsing e Frank Uhlig [95] observaram que o contradomínio numérico em espaços de Krein está intimamente relacionado com

$$K(F, G, H) \;=\; \bigcup_{\beta \geq 0} \beta\, W(F, G, H),$$

o cone gerado por (1.8), quando $F = \mathrm{Re}(HA)$, $G = \mathrm{Im}(HA)$ e $H \in H_n$ é não-singular, em virtude de

$$W_H^{\pm}(A) \;=\; \{x + iy :\ (x, y, \pm 1) \in K(F, G, H)\}. \tag{1.9}$$

Visto que $K(F, G, H)$ é um cone convexo, é óbvio que $W_H^{\pm}(A)$ são conjuntos convexos, mas $W_H(A)$ pode não o ser. Contudo, mostra-se que $W_H(A)$ é *pseudo-convexo*, isto é, dados dois pontos distintos $z, w \in W_H(A)$, ou $W_H(A)$ contém o segmento de recta de extremos z e w, ou $W_H(A)$ contém a recta definida por z e w, excepto o segmento de recta aberto que os une.

[4] Pode-se definir o contradomínio numérico conjunto de m matrizes Hermíticas, $m \in \mathbb{N}$, de modo óbvio. Este conjunto revela-se útil, por exemplo, em teoria de sistemas e estabilidade robusta [51].

Teorema 1.2.1 *Se $A \in M_n$, então $W_H(A)$ é um conjunto pseudo-convexo.*

Demonstração [95]**:** Suponhamos que $W_H(A)$ não é um conjunto singular. A convexidade de $W_H^+(A)$ garante que $W_H(A)$ contém os segmentos de recta que unem todos os pares de pontos distintos de $W_H^+(A)$. Análogo raciocínio vale para pontos de $W_{-H}^+(A)$.

Se H é definida positiva (negativa), então $W_{-H}^+(A)$ ($W_H^+(A)$) é um conjunto vazio e nada mais há a provar, ou seja, verifica-se que $W_H(A)$ é convexo.

Se H é indefinida, então $W_H^+(A)$ e $W_{-H}^+(A)$ são ambos conjuntos não vazios. Sejam $z \in W_H^+(A)$ e $w \in W_{-H}^+(A)$ distintos. Tomam-se $x = \mathrm{Re}z$, $y = \mathrm{Im}z$, $u = \mathrm{Re}w$ e $v = \mathrm{Im}w$. Sendo um cone convexo, $K(F, G, H)$ contém o segmento de recta $\mathcal{L}$ que une $(x, y, 1)$ a $-(u, v, 1)$. Se $\alpha \geq 1$, então $-\alpha^{-1}(u, v, 1) \in \alpha^{-1}\mathcal{L}$ e

$$\alpha(x, y, 1) + (1 - \alpha)(u, v, 1) \ = \ \alpha^2\big(\alpha^{-1}(x, y, 1) + (1 - \alpha^{-1})(-\alpha^{-1})(u, v, 1)\big) \in \bigcup_{\beta>0} \beta\mathcal{L}.$$

Como $K(F, G, H)$ contém $\bigcup_{\beta>0} \beta\mathcal{L}$, usando (1.9), vem $\alpha z + (1 - \alpha)w \in W_H^+(A)$, para todo o $\alpha \geq 1$, e $W_H^+(A)$ é ilimitado. Se $\alpha \leq 0$, pode provar-se, de modo similar, que $\alpha(x, y, 1)+(1-\alpha)(u, v, 1) \in -K(F, G, H)$, o que equivale a ter $\alpha z+(1-\alpha)w \in -W_H^-(A)$, de novo por (1.9). Em síntese,

$$\{\alpha z + (1 - \alpha)w : \ \alpha \leq 0 \ \text{ ou } \ \alpha \geq 1\} \subseteq W_H(A)$$

e, portanto, $W_H(A)$ é ilimitado e pseudo-convexo. ■

Se $H \in H_n$ é definida positiva ou negativa, então $W_H(A)$ é compacto e convexo.

Se $H \in H_n$ não-singular é indefinida, então $W_H(A)$ é um conjunto singular ou um conjunto ilimitado, pseudo-convexo e não necessariamente fechado [95, 97].

1.2.2 Curva Geradora de Fronteira

Chamam-se *pontos angulosos* e *rectas de suporte* de $W_H(A)$ aos pontos angulosos e rectas de suporte, respectivamente, das componentes convexas $W_H^+(A)$ e $W_{-H}^+(A)$. As rectas de suporte podem não existir e, existindo, podem não ser únicas.

Existe uma relação entre os pontos angulosos de $W_H(A)$ e os valores próprios de A, como descrito no lema seguinte, originalmente obtido em [97, Teorema 3.1] e para o qual apresentamos uma nova demonstração na Secção 1.3.3 (vide Nota 1.3.1).

Lema 1.2.1 *Seja $A \in M_n$. Se z é um ponto anguloso do conjunto $\pm W_H^{\pm}(A)$, então $z \in \sigma_H^{\pm}(A)$ e existe $x \in \mathbb{C}^n$, tal que $Ax = zx$, $A^{[*]}x = \bar{z}x$ e $\langle x, x\rangle_H = \pm 1$.*

Dada $A \in M_n$, um subespaço E de $\mathbb{C}^n$ diz-se *A-redutível*, em relação ao produto interno indefinido induzido por H, se $Ax \in E$ e $A^{[*]}x \in E$, para todo o $x \in E$.

Em particular, o Lema 1.2.1 afirma que se z é um ponto anguloso de $W_H(A)$, então $z \in \sigma_H(A)$ e existe um vector próprio anisotrópico-H de A associado a z que gera um subespaço próprio A-redutível.

Francis D. Murnaghan [105] e Rudolf Kippenhahn [86] provaram, independentemente, que a fronteira do contradomínio numérico clássico de uma matriz $A \in M_n$ é gerada pelo conjunto $C(A)$ dos pontos reais da curva algébrica, cuja equação em coordenadas lineares homogéneas é

$$\det\left(u\,\mathrm{Re}A + v\,\mathrm{Im}A + wI_n\right) \;=\; 0. \tag{1.10}$$

(Para um tratamento pormenorizado de curvas algébricas ver, por exemplo, [35, 56, 130]). Esta curva algébrica tem *classe* n, o que significa que por um ponto genérico do plano passam n rectas tangentes à curva. Além disso, os seus n focos reais são os valores próprios da matriz A [53, 121]. O conjunto $W(A)$ é o invólucro convexo da curva $C(A)$ e $\partial W(A)$ é uma união de arcos algébricos (arcos de $C(A)$ e porventura segmentos de recta). Por esta razão, $C(A)$ designa-se a *curva geradora da fronteira* de $W(A)$. Generalizamos este resultado ao contradomínio numérico-H.

Teorema 1.2.2 *Seja $A \in M_n$. Se $ux + vy + w = 0$ é a equação de uma recta de suporte de $W_H(A)$, então*

$$\det\left(u\,\mathrm{Re}^H A + v\,\mathrm{Im}^H A + wI_n\right) \;=\; 0. \tag{1.11}$$

Demonstração: Escreva-se $A_\theta^H = \mathrm{Re}^H\big(\mathrm{e}^{-i\theta}A\big)$, $\theta \in [0, 2\pi)$. Como A_θ^H é uma matriz Hermítica-H, pela propriedade $\mathrm{W_H}7$ e Teorema 1.2.1, $W_H\big(A_\theta^H\big)$ é um subconjunto pseudo--convexo do eixo real. Seja Ω o conjunto dos ângulos $\theta \in [0, 2\pi)$ para os quais $W_H\big(A_\theta^H\big)$ é um segmento de recta fechado ou a união disjunta de duas semi-rectas fechadas de extremos z_1^θ e z_2^θ. Como z_j^θ é um ponto anguloso de $W_H\big(A_\theta^H\big)$, pelo Lema 1.2.1, conclui-se que z_j^θ é um valor próprio de A_θ^H, portanto,

$$\det\left(A_\theta^H - z_j^\theta I_n\right) \;=\; 0, \qquad j = 1, 2. \tag{1.12}$$

Por outro lado, é claro que $\mathcal{L}_j = \left\{z_j^\theta + iy : y \in \mathbb{R}\right\}$, $j = 1, 2$, são rectas de suporte de $W_H\big(\mathrm{e}^{-i\theta}A\big)$. Submetendo as rectas $\mathcal{L}_j$ a uma rotação segundo o ângulo θ, obtêm-se as rectas de suporte de $\mathrm{e}^{i\theta}\, W_H\big(\mathrm{e}^{-i\theta}A\big) = W_H(A)$ de equações

$$\cos\theta\, x + \sin\theta\, y \;=\; z_j^\theta, \qquad j = 1, 2.$$

À medida que θ percorre Ω, descrevem-se todas as rectas de suporte de $W_H(A)$. Assim, qualquer que seja a recta de suporte de $W_H(A)$ de equação $ux + vy + w = 0$, existem um ângulo $\theta_0 \in \Omega$, um escalar não nulo $r \in \mathbb{R}$ e $j \in \{1, 2\}$, tais que

$$u = r\cos\theta_0, \qquad v = r\sin\theta_0 \qquad \text{e} \qquad w = -r\, z_j^{\theta_0}. \tag{1.13}$$

Tomando $\theta = \theta_0$ em (1.12), atendendo a que $A^H_{\theta_0} = \cos\theta_0 \operatorname{Re}^H A + \sin\theta_0 \operatorname{Im}^H A$ e a (1.13), obtém-se $r^{-n} \det\left(u \operatorname{Re}^H A + v \operatorname{Im}^H A + w I_n\right) = 0$, donde vem a equação (1.11). ■

Se $H = \pm I_n$, então o Teorema 1.2.2 reduz-se ao Teorema de Murnaghan-Kippenhahn.

Denota-se o *plano projectivo complexo* por $\mathbb{CP}^2$. Como é usual, identifica-se cada ponto $(x, y) \in \mathbb{C}^2$ com $[x, y, 1] \in \mathbb{CP}^2$, e identifica-se cada ponto $[x, y, z] \in \mathbb{CP}^2$, tal que $z \neq 0$, com $\left(\frac{x}{z}, \frac{y}{z}\right) \in \mathbb{C}^2$.

Nota 1.2.1 Dada $A \in M_n$, como

$$p_A(u, v, w) = \det\left(u \operatorname{Re}^H A + v \operatorname{Im}^H A + w I_n\right)$$

é um polinómio homogéneo de grau n, então o conjunto

$$\Upsilon_{p_A} = \{[u, v, w] \in \mathbb{CP}^2 : \ p_A(u, v, w) = 0\}$$

define uma curva algébrica de grau n. A sua curva dual é dada por

$$\Upsilon^{\wedge}_{p_A} = \{[x, y, z] \in \mathbb{CP}^2 : \ ux + vy + wz = 0 \ \text{ é uma recta tangente a } \Upsilon_{p_A}\}.$$

Denota-se por $C_H(A)$ a parte real desta curva algébrica (de classe n), isto é, o conjunto dos pontos $x + iy$, com $(x, y) \in \mathbb{R}^2$, tais que $ux + vy + w = 0$ é uma recta tangente a Υ_{p_A}, e $C_H(A)$ denomina-se a *curva geradora da fronteira* de $W_H(A)$.

Em geral, uma curva algébrica de classe n tem n^2 focos, incluindo multiplicidades. Se os coeficientes da equação da curva são reais, então exactamente n de entre os focos são reais [121]. Neste caso, o polinómio $p_A(u, v, w)$ tem coeficientes reais e os n focos reais da curva $C_H(A)$ são exactamente os valores próprios da matriz A.

Se $A \in M_2$, então Υ_{p_A} define uma curva algébrica de grau dois, isto é, uma cónica. A sua curva dual é novamente uma cónica. Assim, $C_H(A)$ é uma hipérbole, parábola ou elipse, possivelmente degeneradas:

I. Se A é uma matriz escalar, pela propriedade $\mathrm{W_H}6$, tem-se que $W_H(A)$ é um conjunto singular;

   Caso contrário,

II. Se $H \in H_2$ é definida positiva ou negativa, então $W_H(A)$ é limitado e convexo, consequentemente, $C_H(A)$ é uma elipse, cujos focos são os valores próprios de A, e $W_H(A)$ é o invólucro convexo de $C_H(A)$. Obtém-se, assim, um disco elíptico, cujos eixos se podem explicitar, usando o Teorema do Contradomínio Elíptico, dado que $W_H(A) = W(SAS^{-1})$, onde $S \in M_2$ não-singular é tal que $S^*S = \pm H$. Efectivamente, ao substituir A por SAS^{-1} em (1.5), obtêm-se expressões para os eixos, em tudo análogas às do caso clássico, excepto que $A^{[*]}$ toma o lugar de A^*.

III. Se $H \in H_2$ não-singular é indefinida, então o conjunto $W_H(A)$ é ilimitado e pseudo-convexo, podendo não ser convexo, ou seja, $C_H(A)$ é necessariamente uma hipérbole, cujos focos são os valores próprios de A, e $W_H(A)$ consiste na hipérbole e seu interior. (O Teorema do Contradomínio Hiperbólico que caracteriza completamente este conjunto, incluindo os casos degenerados, será pormenorizado na Subsecção 1.2.4.)

Uma matriz $U \in M_n$ diz-se *pseudo-unitária de assinatura* $(r, n-r)$, $0 \leq r \leq n$, se a correspondente aplicação linear preservar a forma quadrática Hermítica

$$q(x) \;=\; |x_1|^2 + \cdots + |x_r|^2 - |x_{r+1}|^2 - \cdots - |x_n|^2, \qquad x \;=\; (x_1, \ldots, x_n) \;\in\; \mathbb{C}^n,$$

o que equivale a afirmar que $U^*JU = J$ [114], ou seja, U é unitária-J, para $J = I_r \oplus -I_{n-r}$. Denota-se o grupo das matrizes pseudo-unitárias de assinatura $(r, n-r)$ por $U_{r,n-r}$.

Se r é o número de valores próprios positivos da matriz não-singular $H \in H_n$, contando multiplicidades, então $J = I_r \oplus -I_{n-r}$ é a *matriz de inércia* de H. A conhecida *lei da inércia de Sylvester* [80, p.222] assegura a existência de $R \in M_n$ não-singular, tal que $R^*HR = J$. Dada $A \in M_n$, considerando $A_R = R^{-1}AR$, verifica-se facilmente que

$$W_H(A) \;=\; W_J(A_R). \tag{1.14}$$

Além disso, A é uma matriz Hermítica-H (normal-H, unitária-H) se e só se A_R é uma matriz Hermítica-J (normal-J, unitária-J, respectivamente). Portanto, investigar o contradomínio numérico-H de uma das classes-H de matrizes acima mencionadas pode reduzir-se ao estudo do contradomínio numérico-J da respectiva classe-J.

Em particular, se H é definida positiva $(r = n)$ ou negativa $(r = 0)$, então $J = \pm I_n$ e $W_H(A) = W(A_R)$, pelo que se estendem muitas propriedades do contradomínio numérico clássico a $W_H(A)$. Contudo, quando H é indefinida $(0 < r < n)$, ocorrem diferenças entre $W_H(A)$ e a versão clássica, aqui residindo o maior desafio deste estudo.

Assim, doravante e sem perda de generalidade, considera-se $J = I_r \oplus -I_{n-r}$, com $0 \leq r \leq n$, e o *contradomínio numérico-J* de matrizes é o objecto central a investigar.

Lema 1.2.2 *Seja $A \in M_n$ e $x \in \mathbb{C}^n$ tal que $\langle x, x\rangle_J = \pm 1$. Se $W_J^{\pm}\big(\mathrm{Re}^J A\big)$ é um segmento de recta fechado ou uma semi-recta fechada, então as condições seguintes são equivalentes:*

(a) $\mathrm{Re}\,\langle Ax, x\rangle_J$ *é um ponto extremo de* $\pm\mathrm{Re}\,W_J^{\pm}(A)$;

(b) $\langle \mathrm{Re}^J Ax, x\rangle_J$ *é um ponto extremo de* $\pm W_J^{\pm}\big(\mathrm{Re}^J A\big)$;

(c) $\mathrm{Re}^J Ax = \lambda_{\mathrm{M}} x$, *onde* λ_{M} *é o máximo ou o mínimo de* $\sigma_J^{\pm}\big(\mathrm{Re}^J A\big)$.

Demonstração: Dado que $\overline{\langle Ax, x\rangle}_J = \langle A^{[*]}x, x\rangle_J$, então $\mathrm{Re}\,\langle Ax, x\rangle_J = \langle \mathrm{Re}^J Ax, x\rangle_J$ e é óbvia a equivalência entre (a) e (b). Seja z um ponto extremo de $W_J^{+}\big(\mathrm{Re}^J A\big)$. Então z é um ponto anguloso de $W_J^{+}\big(\mathrm{Re}^J A\big)$ e, pelo Lema 1.2.1, z é o valor próprio máximo ou mínimo em $\sigma_J(\mathrm{Re}^J A)$, isto é, $z = \lambda_M$.

(c) $\Rightarrow$ (b) Se x é um vector próprio de $\mathrm{Re}^J A$ associado a λ_M, então $\langle \mathrm{Re}^J Ax, x\rangle_J = \lambda_M$. Logo, $\langle \mathrm{Re}^J Ax, x\rangle_J$ é um extremo de $W_J^{+}\big(\mathrm{Re}^J A\big)$.

(b) $\Rightarrow$ (c) Se x não fosse vector próprio de $\mathrm{Re}^J A$ associado ao valor próprio λ_M, então $\langle \mathrm{Re}^J Ax, x\rangle_J \neq \lambda_M$, e λ_M não seria um ponto extremo de $W_J^{\pm}\big(\mathrm{Re}^J A\big)$, uma contradição.

Análogo raciocínio garante a equivalência, quando z é um ponto extremo da componente convexa $-W_J^{-}\big(\mathrm{Re}^J A\big)$. ■

O conjunto $W_J(A)$ obtém-se da curva geradora de fronteira $C_J(A)$ como o seu *invólucro pseudo-convexo* da seguinte forma. Dados z_1 e z_2 dois pontos arbitrários da curva $C_J(A)$, sejam $x_i \in \mathbb{C}^n$ tais que

$$\frac{\langle A\,x_i, x_i\rangle_J}{\langle x_i, x_i\rangle_J} = z_i, \qquad i = 1, 2;$$

se $\langle x_1, x_1\rangle_J \langle x_2, x_2\rangle_J > 0$, então $W_J(A)$ contém o segmento de recta fechado que une z_1 a z_2; se $\langle x_1, x_1\rangle_J \langle x_2, x_2\rangle_J < 0$, então $W_J(A)$ contém as semi-rectas

$$\{\alpha z_1 + (1-\alpha) z_2 :\ \alpha \leq 0 \ \text{ ou } \ \alpha \geq 1\}.$$

Teorema 1.2.3 *Seja $A \in M_n$ e seja Ω o conjunto dos ângulos $\theta \in [0, 2\pi)$ para os quais os n valores próprios de $\mathrm{Re}^J\big(\mathrm{e}^{-i\theta}A\big)$ são todos reais, cujos vectores próprios associados $u_1^\theta, \ldots, u_n^\theta$ são anisotrópicos-J. Se Ω é um subconjunto não vazio de $[0, 2\pi)$, então $W_J(A)$ é o invólucro pseudo-convexo de*

$$\left\{ \frac{\langle Au_k^\theta, u_k^\theta\rangle_J}{\langle u_k^\theta, u_k^\theta\rangle_J} :\ \theta \in \Omega,\ k = 1, \ldots, n \right\}.$$

Demonstração: Seja $p_A(u, v, w) = \det\left(u\,\mathrm{Re}^J A + v\,\mathrm{Im}^J A + wI_n\right)$. A curva $C_J(A)$ é a parte real da curva $\Upsilon_{p_A}^{\wedge}$, dual da curva algébrica determinada por $p_A(u, v, w) = 0$, no plano projectivo complexo $\mathbb{CP}^2$. No Teorema 1.2.2, mostrou-se que se $ux + vy + w = 0$ é a equação de uma recta de suporte de $W_J(A)$, então $p_A(u, v, w) = 0$. Como a curva dual a $\Upsilon^{\wedge}_{p_A}$ é a curva original definida por $p_A(u,v,w)=0$, inferimos, em particular, que toda a recta de suporte de $W_J(A)$ é tangente a $C_J(A)$.

Consideramos a direcção real $(u,v)=(\cos\theta,\sin\theta)$, $\theta\in\Omega$, para a qual $p_A(u,v,w)=0$ fornece n valores próprios reais $\lambda_1^\theta,\dots,\lambda_n^\theta$ para

$$\mathrm{Re}^J\big(\mathrm{e}^{-i\theta}A\big) \;=\; \cos\theta\,\mathrm{Re}^J A+\sin\theta\,\mathrm{Im}^J A.$$

Por hipótese, os vectores próprios que lhe estão associados $u_1^\theta,\dots,u_n^\theta$, $\theta\in\Omega$, são anisotrópicos-J. Se os valores próprios correspondentes a vectores próprios u_k^θ, satisfazendo $\langle u_k^\theta,u_k^\theta\rangle_J>0$, são todos maiores (ou menores) que os valores próprios correspondentes a vectores próprios u_k^θ, tais que $\langle u_k^\theta,u_k^\theta\rangle_J<0$, então existem rectas de suporte de $W_J(A)$. Seja λ_M^θ o valor próprio máximo (ou mínimo) de $\mathrm{Re}^J\big(\mathrm{e}^{-i\theta}A\big)$, tal que $\cos\theta\,x+\sin\theta\,y=\lambda_M^\theta$ é uma recta de suporte $\mathcal{L}$ de $W_J(A)$. Aplicando o Lema 1.2.2 à matriz $\mathrm{e}^{-i\theta}A$, conclui-se que a intersecção $\mathcal{L}\cap W_J^{\pm}(A)$ consiste em todos os pontos $\langle Au,u\rangle_J$ para os quais u é um vector próprio de $\mathrm{Re}^J\big(\mathrm{e}^{-i\theta}A\big)$ associado a λ_M^θ, tal que $\langle u,u\rangle_J=\pm1$. Estes pontos pertencem a $C_J(A)$. Pela pseudo-convexidade do contradomínio numérico-J, conclui-se que $W_J(A)$ é o invólucro pseudo-convexo dos pontos

$$\frac{\langle Au_k^\theta,u_k^\theta\rangle_J}{\langle u_k^\theta,u_k^\theta\rangle_J},\quad \theta\in\Omega,\quad k=1,\dots,n.$$

Se não existirem rectas de suporte de $W_J(A)$, então em qualquer direcção real de $\mathbb{R}^2$, digamos $(\cos\theta,\sin\theta)$, $\theta\in[0,2\pi)$, existem n rectas tangentes reais a $C_J(A)$ com esta direcção, nomeadamente, determinadas pelos n valores próprios $\lambda_1^\theta,\dots,\lambda_n^\theta$, assim se concluindo a demonstração. ∎

O Teorema 1.2.3 verifica-se, em especial, quando os n valores próprios da matriz $\mathrm{Re}^J\big(\mathrm{e}^{-i\theta}A\big)$ são todos reais e simples. Esta hipótese mais restritiva garante que todos os vectores próprios associados sejam anisotrópicos-J.

1.2.3 Resultados para Matrizes Hermíticas-J e Normais-J

Seja $J=I_r\oplus -I_{n-r}$, $0\le r\le n$. Como já referido em $\mathrm{W_H}7$, se A é uma matriz Hermítica-J, então $W_J(A)$ é um subconjunto do eixo real. Provamos de seguida que, além disso, vale a igualdade $W_J(A)=\mathbb{R}$, se A admitir pelo menos um valor próprio não real.

Teorema 1.2.4 *Se $A\in M_n$ é Hermítica-J e se os seus valores próprios não são todos reais, então $W_J(A)=\mathbb{R}$.*

Demonstração: Suponhamos que os valores próprios da matriz A são $\lambda_1,\bar\lambda_1,\dots,\lambda_s,\bar\lambda_s$, $\lambda_{2s+1},\dots,\lambda_n$, todos simples, e que apenas $\lambda_{2s+1},\dots,\lambda_n$ são reais. Então existe uma base

$\mathcal{B}_A$ de vectores próprios associados $f_1, g_1, \ldots, f_s, g_s, f_{2s+1}, \ldots, f_n$ satisfazendo

$$\langle f_k, f_j \rangle_J \;=\; \langle g_k, g_j \rangle_J \;=\; \langle f_k, g_j \rangle_J \;=\; 0, \qquad k \neq j, \quad k, j = 1, \ldots, n, \tag{1.15}$$

$$\langle f_k, f_k \rangle_J \;=\; \langle g_k, g_k \rangle_J \;=\; 0, \qquad k = 1, \ldots, s. \tag{1.16}$$

De facto, se $\langle f_k, f_j \rangle_J \neq 0$, como A é Hermítica-J, então

$$\lambda_k \;=\; \frac{\langle Af_k, f_j \rangle_J}{\langle f_k, f_j \rangle_J} \;=\; \frac{\langle f_k, Af_j \rangle_J}{\langle f_k, f_j \rangle_J} \;=\; \bar{\lambda}_j, \qquad k \neq j, \quad k, j = 1, \ldots, n,$$

o que contradiz o facto de os valores próprios serem todos distintos. De modo análogo, se obtêm as restantes relações em (1.15) e (1.16).

Seja D a matriz em M_n cujas colunas são os vectores próprios da base $\mathcal{B}_A$. Atendendo a (1.15), a matriz D^*JD é diagonal por blocos 2×2 e 1×1, correspondentes aos valores próprios complexos e reais, respectivamente. Atendendo a (1.16), os blocos 2×2 têm as entradas diagonais nulas e, como $\det(D^*JD) \neq 0$, as entradas fora da diagonal são não nulas. Portanto, $\gamma_k = \langle g_k, f_k \rangle_J \neq 0$ e, consequentemente, $\langle Ag_k, f_k \rangle_J = \bar{\lambda}_k \gamma_k$, $k = 1, \ldots, s$.

Considere-se o subconjunto $R_k(A)$ de $W_J(A)$ definido por

$$R_k(A) \;=\; \left\{ \frac{\langle Av_k^\theta, v_k^\theta \rangle_J}{\langle v_k^\theta, v_k^\theta \rangle_J} : \;\; v_k^\theta = f_k + \mathrm{e}^{i\theta} g_k, \quad \theta \in [0, 2\pi), \quad \langle v_k^\theta, v_k^\theta \rangle_J \neq 0 \right\}.$$

Mediante alguns cálculos, prova-se que os elementos de $R_k(A)$, $k = 1, \ldots, s$, são da forma

$$\frac{\langle Av_k^\theta, v_k^\theta \rangle_J}{\langle v_k^\theta, v_k^\theta \rangle_J} \;=\; \frac{\mathrm{Re}(\mathrm{e}^{i\theta} \bar{\lambda}_k \gamma_k)}{\mathrm{Re}(\mathrm{e}^{i\theta} \gamma_k)} \;=\; |\lambda_k| \left(\cos \xi_k + \sin \xi_k \tan \phi_k \right),$$

com $\xi_k = \arg \lambda_k \neq 0$ e $\phi_k = \theta + \arg \gamma_k \neq \pi/2 + n\pi$, $n \in \mathbb{Z}$. Como $\lambda_k \neq 0$, $k = 1, \ldots, s$, à medida que θ varia de 0 a 2π, os elementos de $R_k(A)$ percorrem todo o eixo real. Sendo A Hermítica-J, então $R_k(A) \subseteq W_J(A) \subseteq \mathbb{R}$. Portanto, $W_J(A) = \mathbb{R}$.

Se os valores próprios de A não forem todos simples, por uma perturbação, podemos torná-los todos distintos e o resultado segue por um argumento de continuidade. ■

Como constataremos na Subsecção 1.2.4, o recíproco do Teorema 1.2.4 vale para matrizes de dimensão dois. Contudo, para matrizes de dimensão superior, $W_J(A)$ pode ser todo o eixo real, mesmo que os valores próprios da matriz A Hermítica-J sejam todos reais. Tal acontece, por exemplo, quando $A = \mathrm{diag}\,(1, 3, 2)$ e $J = \mathrm{diag}\,(1, 1, -1)$, pelo facto de $1, 3 \in \sigma_J^+(A)$, $2 \in \sigma_J^-(A)$ e atendendo à pseudo-convexidade do contradomínio numérico-J.

Uma matriz $A \in M_n$ diz-se *essencialmente Hermítica-J* se existirem $\alpha, \beta \in \mathbb{C}$, $\beta \neq 0$, tais que $\alpha I_n + \beta A$ é Hermítica-J.

Corolário 1.2.1 *Se $A \in M_n$ é essencialmente Hermítica-J e se os valores próprios da matriz Hermítica-J $\alpha I_n + \beta A$, $\alpha, \beta \in \mathbb{C}$, $\beta \neq 0$, não são todos reais, então $W_J(A)$ é a recta que passa por $-\alpha/\beta$ e que tem a direcção de $-\arg\beta$.*

Demonstração: Pelo Teorema 1.2.4, tem-se $W_J(\alpha I_n + \beta A) = \mathbb{R}$. Pela propriedade $\mathrm{W_H}2$, vale $W_J(\alpha I_n + \beta A) = \alpha + \beta\, W_J(A)$ e o resultado segue facilmente. ■

O próximo corolário é especialmente útil, pois relaciona os valores próprios da matriz $\mathrm{Re}^J\left(\mathrm{e}^{i\theta}A\right)$ com a existência de uma recta de suporte de $W_J(A)$ segundo um determinado ângulo perpendicular a θ.

Corolário 1.2.2 *Seja $A \in M_n$ e $\theta \in [0, 2\pi)$. Se $W_J(A)$ tem uma recta de suporte na direcção perpendicular à do ângulo θ, então os valores próprios de $\mathrm{Re}^J\left(\mathrm{e}^{i\theta}A\right)$ são todos reais.*

Demonstração: Se a matriz Hermítica-J $\mathrm{Re}^J\left(\mathrm{e}^{i\theta}A\right)$ admitisse pelo menos um valor próprio não real, pelo Teorema 1.2.4, ter-se-ia $W_J\left(\mathrm{Re}^J\left(\mathrm{e}^{i\theta}A\right)\right) = \mathbb{R}$. Como as propriedades $\mathrm{W_H}2$ e $\mathrm{W_H}8$ garantem que $\mathrm{Re}\left(\mathrm{e}^{i\theta}W_J(A)\right) = W_J\left(\mathrm{Re}^J\left(\mathrm{e}^{i\theta}A\right)\right)$, conclui-se que a projecção ortogonal da rotação de $W_J(A)$ segundo o ângulo θ seria todo o eixo real. Portanto, $W_J(A)$ não teria nenhuma recta de suporte na direcção perpendicular à do ângulo θ. ■

Apesar do recíproco do Corolário 1.2.2 se verificar no caso 2×2, não é válido em geral.

Exemplo 1.2.1 Se $J = \mathrm{diag}\,(1, -1, -1)$ e

$$A \;=\; \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix},$$

então os valores próprios de $\mathrm{Re}^J\left(\mathrm{e}^{i\theta}A\right)$, $\theta \in [0, 2\pi)$, são os números reais $\cos\theta$ e $\cos\theta \pm 1$, contudo, $W_J(A)$ não possui recta de suporte em direcção alguma. Efectivamente, a curva geradora da fronteira $C_J(A)$ é constituída pelo ponto $(1, 0)$ e pela circunferência de centro $(1, 0)$ e raio de comprimento um; cujo invólucro pseudo-convexo é todo o plano complexo.

Chi-Kwong Li e Leiba Rodman [96] implementaram um programa em Matlab para gerar uma aproximação ao conjunto $W_H^+(A)$, a partir do contradomínio numérico conjunto de um triplo de matrizes Hermíticas. Seguindo uma abordagem alternativa, na linha do Teorema 1.2.3 e dos resultados obtidos na presente subsecção, desenvolvemos um algoritmo, com vista a gerar computacionalmente uma aproximação ao conjunto $W_J(A)$ e suas curvas geradoras de fronteira (consultar [22]). A ideia do algoritmo é baseada na existência ou não de rectas de suporte em diferentes direcções.

Como já referimos, a propósito da propriedade W1', o contradomínio numérico clássico de uma matriz normal $A \in M_n$ é o invólucro convexo do espectro de A. A validade de um resultado análogo para $W_J(A)$, quando A é uma matriz normal-J, é uma questão levantada em [95]. Respondemos afirmativamente a esta questão, quando A tem valores próprios simples.

Teorema 1.2.5 *Se $A \in M_n$ é uma matriz normal-J de valores próprios simples, então $W_J(A)$ é o invólucro pseudo-convexo de $\sigma(A)$.*

Demonstração: Como os valores próprios $\lambda_1, \ldots, \lambda_n$ de A são todos distintos, existe uma base de vectores próprios $v_1, \ldots, v_n$ satisfazendo $Av_k = \lambda_k v_k$ e $\langle v_k, v_k \rangle_J \neq 0$, $k = 1, \ldots, n$. Por hipótese, A comuta com $A^{[*]}$, a sua adjunta-J, então $AA^{[*]}v_k = \lambda_k A^{[*]}v_k$, isto é, $w_k = A^{[*]}v_k$ é um vector próprio de A associado a λ_k. A independência linear de $v_1, \ldots, v_n$ assegura que exista $c_k \in \mathbb{C}$ tal que $w_k = c_k v_k$. Atendendo a $A^{[*]}v_k = c_k v_k$ e $Av_k = \lambda_k v_k$, tem-se

$$c_k \langle v_k, v_k \rangle_J \;=\; \langle A^{[*]}v_k, v_k \rangle_J \;=\; \langle v_k, Av_k \rangle_J \;=\; \bar{\lambda}_k \langle v_k, v_k \rangle_J,$$

logo $c_k = \bar{\lambda}_k$. Mediante cálculos básicos, obtém-se

$$\mathrm{Re}^J\left(\mathrm{e}^{-i\theta}A\right) v_k \;=\; \mathrm{Re}\left(\mathrm{e}^{-i\theta}\lambda_k\right) v_k, \qquad k = 1, \ldots, n, \qquad \theta \in [0, 2\pi),$$

o que significa que $v_1, \ldots, v_n$ são vectores próprios de $\mathrm{Re}^J\left(\mathrm{e}^{-i\theta}A\right)$ independentes de θ e que os valores próprios de $\mathrm{Re}^J\left(\mathrm{e}^{-i\theta}A\right)$ são todos reais. Nestas circunstâncias, podemos aplicar o Teorema 1.2.3, considerando $u_k^\theta = v_k$, qualquer que seja θ. Da igualdade

$$\sigma(A) \;=\; \left\{ \frac{\langle Av_k, v_k \rangle_J}{\langle v_k, v_k \rangle_J} : \; k = 1, \ldots, n \right\},$$

inferimos que $W_J(A)$ é o invólucro pseudo-convexo do espectro de A. ■

A existência de valores próprios múltiplos de A pode originar a existência de vectores próprios isotrópicos-J. Por exemplo, se $J = \mathrm{diag}\,(1, -1)$, então

$$A \;=\; \begin{bmatrix} 3 & -1 \\ 1 & 1 \end{bmatrix}$$

é uma matriz Hermítica-J de valor próprio duplo 2 e os vectores próprios que lhe estão associados são isotrópicos-J. Ora, como constataremos a partir do Teorema 1.2.7 e) ii, $W_J(A) = \mathbb{R} \setminus \{2\}$, não sendo aqui aplicável o conceito de invólucro pseudo-convexo do espectro de A.

É bem conhecido [81] que $A \in M_n$ é uma matriz unitária se e só se $W(A)$ é um polígono inscrito no disco unitário $\mathcal{D}$ e $\sigma(A) \subseteq \partial\mathcal{D}$. A validade de um resultado análogo para $W_J(U)$, quando U é uma matriz pseudo-unitária, é outra questão levantada em [95]. Neste sentido, apresenta-se o seguinte resultado.

Corolário 1.2.3 *Se $U \in U_{r,n-r}$ tem valores próprios simples, então $W_J(U)$ é o invólucro pseudo-convexo de $\sigma(U) \subseteq \partial\mathcal{D}$.*

Demonstração: A simplicidade dos valores próprios $\lambda_1, \ldots, \lambda_n$ de U assegura a existência de uma base de vectores próprios anisotrópicos-J de U, digamos $v_1, \ldots, v_n$ que lhes estão associados. Como $U^{[*]}U = I_n$ e $Uv_k = \lambda_k v_k$, tem-se $\lambda_k U^{[*]} v_k = v_k$ e, portanto,

$$\langle v_k, v_k\rangle_J \;=\; \lambda_k \langle U^{[*]} v_k, v_k\rangle_J \;=\; \lambda_k \langle v_k, U v_k\rangle_J \;=\; \lambda_k \bar{\lambda}_k \langle v_k, v_k\rangle_J, \tag{1.17}$$

concluindo-se que $|\lambda_k| = 1$, $k = 1, \ldots, n$, ou seja, $\sigma(U) \subseteq \partial\mathcal{D}$. Aplicando o Teorema 1.2.5 à matriz U normal-J, prova-se o corolário. ■

1.2.4 Teorema do Contradomínio Hiperbólico

Nesta subsecção, obtém-se um importante teorema que, paralelamente ao Teorema do Contradomínio Elíptico, permite caracterizar, de forma minuciosa, o contradomínio numérico-H de matrizes $A \in M_2$, qualquer que seja $H \in H_2$ não-singular.

Numa primeira abordagem ao problema, Chi-Kwong Li e Leiba Rodman [96] sugerem a forma hiperbólica de $W_H(A)$, quando $H \in H_2$ não-singular é indefinida, e a possibilidade de $W_H^+(A)$ degenerar num subconjunto de uma recta, num semi-plano ou em todo o plano complexo. Contudo, não pormenorizam a caracterização dos conjuntos, nem especificam em que condições concretas cada caso degenerado ocorre. Propomo-nos, seguindo uma via alternativa, obter uma descrição completa. Começamos por um resultado parcial, envolvendo uma classe específica de matrizes.

Teorema 1.2.6 *Sejam $J = \mathrm{diag}\,(1, -1)$,*

$$A_1 \;=\; \begin{bmatrix} 1 & a_1 \mathrm{e}^{i\theta_1} \\ b_1 \mathrm{e}^{i\theta_1} & -1 \end{bmatrix}, \qquad a_1, b_1 \geq 0, \quad a_1 \neq b_1, \quad \theta_1 \in [0, 2\pi),$$

e $N_1 = 2\,|\det(A_1)| + 2 - a_1^2 - b_1^2$.

(i) *Se $N_1 < 0$, então $W_J(A_1)$ é todo o plano complexo.*

(ii) *Se $N_1 = 0$, então $W_J(A_1)$ é todo o plano complexo, excepto uma recta.*

(iii) *Se $N_1 > 0$, então $W_J(A_1)$ é limitado por uma hipérbole centrada na origem e de eixo tranverso de comprimento $\sqrt{N_1}$.*

Demonstração: Considera-se a função

$$\mu(\varphi) \ = \ \frac{1}{4}\left(2 - a_1^2 - b_1^2 + 2\cos(2\varphi) + 2a_1 b_1 \cos(2\theta_1 + 2\varphi)\right), \qquad \varphi \in \mathbb{R}.$$

Os valores próprios da matriz $\mathrm{Re}^J(\mathrm{e}^{i\varphi}A_1)$ são da forma $\lambda_k^\varphi = \pm\sqrt{\mu(\varphi)}$, $k = 1, 2$.

a) Se $\det(A_1) = 0$, isto é, $a_1 b_1 = 1$ e $\theta_1 = \pi/2 + n\pi$, $n \in \mathbb{Z}$, então

$$N_1 \ = \ -\frac{(1 - a_1^2)^2}{a_1^2} \ < \ 0$$

e os valores próprios $\lambda_k^\varphi = \pm\sqrt{N_1}/2$, $i = 1, 2$, são números complexos imaginários puros, independentes de φ. O Corolário 1.2.2 garante, nestas circunstâncias, que $W_J(A_1)$ não tem rectas de suporte em direcção alguma, o que significa que $W_J(A_1) = \mathbb{C}$.

b) Se $\det(A_1) \neq 0$, isto é, $a_1 b_1 \neq 1$ ou $\theta_1 \neq \pi/2 + n\pi$, $n \in \mathbb{Z}$, então μ é uma função real, satisfazendo

$$4\mu(\varphi) \ = \ 2 - a_1^2 - b_1^2 + 2\left|1 + a_1 b_1\,\mathrm{e}^{2i\theta_1}\right|\cos(\delta_1 + 2\varphi) \ \leq \ N_1, \qquad \varphi \in \mathbb{R},$$

onde

$$\tan\delta_1 \ = \ \frac{a_1 b_1 \sin(2\theta_1)}{1 + a_1 b_1 \cos(2\theta_1)}.$$

Assim, a função μ atinge um máximo, dado por $N_1/4$, no ponto $\varphi_0 = n\pi - \delta_1/2$, $n \in \mathbb{Z}$.

(i) Se $N_1 < 0$, é claro que $\mu(\varphi) < 0$ e os valores próprios λ_1^φ e λ_2^φ são números complexos imaginários puros, para todo o $\varphi \in \mathbb{R}$. Como no caso a), conclui-se que $W_J(A_1)$ é todo o plano complexo.

(ii) Se $N_1 = 0$, então λ_1^φ e λ_2^φ são números complexos imaginários puros, qualquer que seja $\varphi \neq \varphi_0$. Pelo Corolário 1.2.2, $W_J(A_1)$ só pode, eventualmente, admitir uma recta de suporte na direcção perpendicular a φ_0. Ora, $\lambda_1^{\varphi_0} = \lambda_2^{\varphi_0} = 0$ e os vectores próprios de A_1 que lhe estão associados são isotrópicos-J. Portanto, $W_J(A_1)$ é todo o plano complexo, excepto a recta $\cos\varphi_0\, x + \sin\varphi_0\, y = 0$.

(iii) Se $N_1 > 0$, então existem intervalos (ξ, η) contendo φ_0, com $\mu(\xi) = \mu(\eta) = 0$, tais que $\mu(\varphi) > 0$, para qualquer $\varphi \in (\xi, \eta)$. Em geral, dado $\varphi \in \mathbb{R}$, os vectores

$$u_k^\varphi \ = \ \left(-2\cos\varphi - 2\lambda_k^\varphi, a_1\mathrm{e}^{-i(\theta_1+\varphi)} + b_1\mathrm{e}^{i(\theta_1+\varphi)}\right)$$

são vectores próprios de $\mathrm{Re}^J(\mathrm{e}^{i\varphi}A_1)$ associados a λ_k^φ que satisfazem

$$\langle u_k^\varphi, u_k^\varphi\rangle_J \ = \ 4\,(\overline{\lambda_k^\varphi} + \cos\varphi)(\lambda_k^\varphi + \cos\varphi) - a_1^2 - b_1^2 + 2a_1 b_1\cos(2\theta_1 + 2\varphi). \qquad (1.18)$$

Dado $\varphi \in (\xi, \eta)$, como $a_1 \neq b_1$, temos

$$0 \ < \ \mu(\varphi) \leq \cos^2\varphi - \frac{1}{4}(a_1 - b_1)^2 \ < \ \cos^2\varphi,$$

portanto, os valores próprios λ_k^φ, $k=1,2$, são números reais não nulos de vectores próprios anisotrópicos-J, já que $\langle u_k^\varphi, u_k^\varphi\rangle_J = 8\lambda_k^\varphi\,(\lambda_k^\varphi + \cos\varphi) \neq 0$, $k=1,2$. Pelo Teorema 1.2.3, os pontos da forma

$$x^\varphi + iy^\varphi \;=\; \frac{\langle A_1 u_k^\varphi, u_k^\varphi\rangle_J}{\langle u_k^\varphi, u_k^\varphi\rangle_J}, \qquad \varphi \in (\xi,\eta)\,,$$

pertencem à curva geradora de fronteira de $W_J(A_1)$, verificando-se que

$$x^\varphi \;=\; \frac{1}{4\lambda_k^\varphi}\left(2a_1b_1\cos(2\theta_1+\varphi) - (a_1^2+b_1^2-4)\cos\varphi\right), \tag{1.19}$$

$$y^\varphi \;=\; \frac{1}{4\lambda_k^\varphi}\left(2a_1b_1\sin(2\theta_1+\varphi) + (a_1^2+b_1^2)\sin\varphi\right). \tag{1.20}$$

Mediante alguns cálculos, a partir de (1.19) e (1.20), vem

$$\begin{cases} \alpha_1\cos(2\varphi) + \beta_1\sin(2\varphi) \;=\; \gamma_1 \\ \alpha_2\cos(2\varphi) + \beta_2\sin(2\varphi) \;=\; \gamma_2 \end{cases}, \tag{1.21}$$

onde

$$\begin{aligned} \alpha_1 &= \mathrm{C}^2 - \mathrm{B}^2 - 8\mathrm{D}x^2, & \beta_1 &= 2\mathrm{B}(\mathrm{C}+4x^2), & \gamma_1 &= 8\mathrm{E}x^2 - \mathrm{B}^2 - \mathrm{C}^2,\\ \alpha_2 &= \mathrm{B}^2 - \mathrm{A}^2 - 8\mathrm{D}y^2, & \beta_2 &= 2\mathrm{B}(\mathrm{A}+4y^2), & \gamma_2 &= 8\mathrm{E}y^2 - \mathrm{A}^2 - \mathrm{B}^2, \end{aligned}$$

e

$$\begin{aligned} \mathrm{A} &= a_1^2+b_1^2+2a_1b_1\cos(2\theta_1), & \mathrm{B} &= 2a_1b_1\sin(2\theta_1),\\ \mathrm{C} &= a_1^2+b_1^2-2a_1b_1\cos(2\theta_1)-4, & \mathrm{D} &= 2+2a_1b_1\cos(2\theta_1),\\ \mathrm{E} &= 2-a_1^2-b_1^2, & \mathrm{F} &= \tfrac{1}{4}\left(a_1^2-b_1^2\right)^2 - \mathrm{A}. \end{aligned}$$

Eliminando φ no sistema (1.21), obtém-se $\mathrm{A}x^2+2\mathrm{B}xy+\mathrm{C}y^2+\mathrm{F}=0$, a equação quadrática de uma cónica que, reduzida aos eixos principais, mediante uma mudança de coordenadas adequada, tem equação

$$\frac{X^2}{N_1} - \frac{Y^2}{M_1} \;=\; \frac{1}{4}, \tag{1.22}$$

com $M_1 = 2\,|\det(A_1)| - 2 + a_1^2 + b_1^2$. Como $M_1 > 0$, a equação (1.22) da curva geradora de fronteira de $W_J(A_1)$ é uma hipérbole centrada na origem, de eixos transverso e não-transverso de comprimento $\sqrt{N_1}$ e $\sqrt{M_1}$, respectivamente. ∎

O próximo lema permitirá simplificar consideravelmente o estudo a que nos propomos.

Lema 1.2.3 *Dada $A=[a_{ij}]\in M_2$, existe uma matriz pseudo-unitária $U\in U_{1,1}$ tal que*

$$U^{-1}AU \;=\; \begin{bmatrix} a_{11} & |a_{12}|\,\mathrm{e}^{i\vartheta} \\ |a_{21}|\,\mathrm{e}^{i\vartheta} & a_{22} \end{bmatrix}, \qquad \vartheta \;=\; \frac{1}{2}\arg\left(a_{12}a_{21}\right).$$

Demonstração: Seja $U = \mathrm{diag}\,(\mathrm{e}^{i\eta}, \mathrm{e}^{i\mu})$, tal que $2\eta + \arg a_{21} = 2\mu + \arg a_{12}$. A matriz U pertence ao grupo pseudo-unitário $U_{1,1}$ e satisfaz o pretendido. ■

É chegado o momento de apresentar, na sua forma mais geral, o Teorema do Contradomínio Hiperbólico, um pilar essencial desta teoria de contradomínios numéricos em espaços de Krein e que se revelará de muito interesse e utilidade não só nas secções seguintes, como ainda na obtenção dos resultados centrais do capítulo dois.

Teorema 1.2.7 (Teorema do Contradomínio Hiperbólico) *Seja* $J = \mathrm{diag}\,(1, -1)$ *e* $A = [a_{ij}] \in M_2$, *tal que* $\sigma(A) = \{\alpha_1, \alpha_2\}$, $\alpha = \frac{1}{2}\mathrm{Tr}(A)$. *Sejam*

$$M \;=\; |\alpha_1|^2 + |\alpha_2|^2 - \mathrm{Tr}\big(A^{[*]}A\big) \qquad e \qquad N \;=\; \mathrm{Tr}\big(A^{[*]}A\big) - 2\,\mathrm{Re}\,(\overline{\alpha}_1\alpha_2)\,. \tag{1.23}$$

a) *Se* $M > 0$ *e* $N > 0$, *então* $W_J(A)$ *é limitado por uma hipérbole de focos* α_1 *e* α_2, *eixos transverso e não-transverso de comprimento* $\sqrt{N}$ *e* $\sqrt{M}$, *respectivamente.*

b) *Se* $M > 0$ *e* $N = 0$, *então* $W_J(A)$ *é:*

i. *a recta* l, *se* $|a_{12}| = |a_{21}|$, **ii.** *todo o plano, excepto a recta* l, *se* $|a_{12}| \neq |a_{21}|$,

em que l *é a recta que passa por* α *e é perpendicular ao segmento que une* α_1 *a* α_2*;*

c) *Se* $M > 0$ *e* $N < 0$, *então* $W_J(A)$ *é todo o plano complexo.*

d) *Se* $M = 0$ *e* $N > 0$, *então* $W_J(A)$ *é a recta definida por* α_1 *e* α_2, *excepto o segmento de recta aberto de extremos* α_1 *e* α_2.

e) *Se* $M = N = 0$, *então* $W_J(A)$ *é:*

i. $\{\alpha\}$, *se* $a_{11} = a_{22}$*;* **ii.** *a recta definida por* a_{11} *e* a_{22}, *excepto* α, *se* $a_{11} \neq a_{22}$.

Demonstração: Consideram-se

$$A_k \;=\; \begin{bmatrix} k & a_k\mathrm{e}^{i\,\theta_k} \\ b_k\mathrm{e}^{i\,\theta_k} & -k \end{bmatrix} \qquad \mathrm{e} \qquad \beta_k \;=\; \frac{1}{2}(a_{11} - a_{22})k + 1 - k,$$

com $k = 0$, se $a_{11} = a_{22}$, e $k = 1$, caso contrário. Pelo Lema 1.2.3, existe uma matriz pseudo-unitária $U \in U_{1,1}$, tal que

$$U^{[*]}AU \;=\; \alpha I_2 + \beta_k A_k, \tag{1.24}$$

em que $U^{[*]}$ é a adjunta-J de U e

$$a_k \;=\; |a_{12}||\beta_k|^{-1}, \qquad b_k \;=\; |a_{21}||\beta_k|^{-1}, \qquad \theta_k \;=\; \frac{1}{2}\arg(a_{12}a_{21}) - \arg\beta_k.$$

Utilizando as propriedades $\mathrm{W_H}1$ e $\mathrm{W_H}2$, podemos descrever $W_J(A)$ a partir de

$$W_J(A_k) \;=\; \left\{ \frac{k|x_1|^2+ k|x_2|^2+ (a_k\overline{x_1}x_2- b_k x_1\overline{x_2})\mathrm{e}^{i\theta_k}}{|x_1|^2- |x_2|^2} \;:\; x_1, x_2 \in \mathbb{C},\; |x_1| \neq |x_2| \right\}, \tag{1.25}$$

com $k = 0$, se $a_{11} = a_{22}$, e $k = 1$, caso contrário. Tomando

$$r \;=\; \frac{|x_1|^2 + |x_2|^2}{|x_1|^2 - |x_2|^2} \qquad \text{e} \qquad \phi \;=\; \arg x_2 - \arg x_1,$$

reescreve-se (1.25) na forma

$$W_J(A_k) \;=\; \left\{ rk + \frac{1}{2}\sqrt{r^2-1}\left(a_k\mathrm{e}^{i\phi} - b_k\mathrm{e}^{-i\phi}\right)\mathrm{e}^{i\theta_k} \;:\; r \in \mathbb{R},\;\; r^2-1 \geq 0,\;\; \phi \in [0, 2\pi) \right\}.$$

Portanto, $x + iy \in W_J(A_k)$ se e só se

$$\begin{aligned} x &= k\,r + \frac{1}{2}\sqrt{r^2-1}\,((a_k - b_k)\,\cos\theta_k\cos\phi - (a_k + b_k)\sin\theta_k\sin\phi), \\ y &= \frac{1}{2}\sqrt{r^2-1}\,((a_k - b_k)\,\sin\theta_k\cos\phi + (a_k + b_k)\cos\theta_k\sin\phi). \end{aligned}$$

I. Seja $k = 1$ e $a_1 \neq b_1$. Mediante alguns cálculos, obtém-se a seguinte família de curvas:

$$\frac{\left((x-r)\sin\theta_1 - y\cos\theta_1\right)^2}{(a_1 + b_1)^2} + \frac{\left((x-r)\cos\theta_1 + y\sin\theta_1\right)^2}{(a_1 - b_1)^2} \;=\; \frac{r^2-1}{4},$$

determinada pelo parâmetro r tal que $r^2 - 1 \geq 0$. A envolvente a esta família de curvas é dada pela equação $\mathrm{A}x^2 + 2\mathrm{B}xy + \mathrm{C}y^2 + \mathrm{F} = 0$ de uma cónica, em que

$$\begin{aligned} \mathrm{A} &= a_1^2 + b_1^2 + 2a_1b_1\cos(2\theta_1), & \mathrm{B} &= 2a_1b_1\sin(2\theta_1), \\ \mathrm{C} &= a_1^2 + b_1^2 - 2a_1b_1\cos(2\theta_1) - 4, & \mathrm{F} &= \tfrac{1}{4}\left(a_1^2 - b_1^2\right)^2 - \mathrm{A}. \end{aligned} \tag{1.26}$$

Reduzindo a cónica aos seus eixos principais, através de uma mudança de coordenadas adequada, obtém-se a equação reduzida

$$M_1X^2 - N_1Y^2 \;=\; \frac{M_1N_1}{4}, \tag{1.27}$$

em que M_1 e $-N_1$ são os valores próprios da matriz simétrica real

$$S \;=\; \begin{bmatrix} \mathrm{A} & \mathrm{B} \\ \mathrm{B} & \mathrm{C} \end{bmatrix}, \tag{1.28}$$

associada à forma quadrática $\mathrm{A}x^2 + 2\mathrm{B}xy + \mathrm{C}y^2$, ou seja,

$$M_1 \;=\; 2\sqrt{1 + 2a_1b_1\cos(2\theta_1) + a_1^2b_1^2} - 2 + a_1^2 + b_1^2, \tag{1.29}$$

$$N_1 \;=\; 2\sqrt{1 + 2a_1b_1\cos(2\theta_1) + a_1^2b_1^2} + 2 - a_1^2 - b_1^2. \tag{1.30}$$

Atendendo a $a_1^2 + b_1^2 \geq 2a_1b_1$, tem-se

$$M_1 \;\geq\; 2|1 - a_1b_1| - 2(1 - a_1b_1) \;\geq\; 0.$$

(Mais, $M_1 = 0$ se e só se $a_1 = b_1 = 0$ ou $a_1 = b_1 \leq 1$ e $\theta_1 = \pi/2 + n\pi$, $n \in \mathbb{Z}$.) Sendo $a_1 \neq b_1$, obviamente $M_1 > 0$. Verifica-se, ainda, facilmente que $\mathrm{Tr}\,(A_1^{[*]}A_1) = 2 - a_1^2 - b_1^2$ e, como os valores próprios de A_1 são $\alpha_1' = \sqrt{1 + a_1b_1\mathrm{e}^{2i\theta_1}}$ e $-\alpha_1'$, tem-se

$$M_1 \;=\; 2|\alpha_1'|^2 - \mathrm{Tr}\,(A_1^{[*]}A_1) \qquad \text{e} \qquad N_1 \;=\; 2|\alpha_1'|^2 + \mathrm{Tr}\,(A_1^{[*]}A_1). \tag{1.31}$$

Por outro lado, os vectores próprios da matriz S em (1.28) associados ao valor próprio M_1 têm a direcção do vector

$$u \;=\; \left(\mathrm{Re}\,(\alpha_1')^2 - |\alpha_1'|^2, \mathrm{Im}\,(\alpha_1')^2\right).$$

Ora, identificando vectores em $\mathbb{R}^2$ com números complexos, u^2 é dado por

$$\left(\mathrm{Re}\,(\alpha_1')^2 - |\alpha_1'|^2 + i\,\mathrm{Im}\,(\alpha_1')^2\right)^2 \;=\; 2\left(\mathrm{Re}\,(\alpha_1')^2 - |\alpha_1'|^2\right)(\alpha_1')^2,$$

concluindo-se que a recta definida pelos valores próprios $-\alpha_1'$ e α_1' tem a direcção do vector u. Podemos, agora, classificar a cónica que descreve a fronteira de $W_J(A_1)$, atendendo ao sinal de M_1N_1, apenas dependente do sinal de N_1, dada a positividade estrita de M_1. Três subcasos podem ocorrer:

a) Se $N_1 > 0$, então (1.27) reduz-se à equação

$$\frac{X^2}{N_1} - \frac{Y^2}{M_1} \;=\; \frac{1}{4},$$

de uma hipérbole centrada na origem, cujos eixos transverso e não-transverso têm comprimento $\sqrt{N_1}$ e $\sqrt{M_1}$, respectivamente, sendo a semi-distância focal dada por $\gamma = |\alpha_1'|$. A direcção do eixo transverso da hipérbole é a do vector u, concluindo-se que os seus focos são $\pm\gamma\, u/\|u\|$, precisamente, os valores próprios de A_1.

b) Se $N_1 = 0$, então (1.27) degenera na equação $X = 0$, da recta l_1 que passa pela origem e é perpendicular ao segmento que une $-\alpha_1'$ a α_1'. Pelo Teorema 1.2.6 (ii), $W_J(A_1)$ é todo o plano complexo, excepto a recta l_1.

c) Se $N_1 < 0$, pelo Teorem 1.2.6 (i), tem-se que $W_J(A_1)$ é todo o plano complexo.

II. Seja $k = 1$ e $a_1 = b_1$.

a) Se $a_1 = 0$ ou $\theta_1 = \pi/2 + n\pi$, $n \in \mathbb{Z}$, então os valores próprios da matriz A_1 reduzem-se a $\alpha_1' = \sqrt{1 - a_1^2}$ e $-\alpha_1'$, e $W_J(A_1)$ é um subconjunto pseudo-convexo de $\mathbb{R}$. Os elementos das componentes convexas $W_J^+(A_1)$ e $-W_J^-(A_1)$ são determinados pelas imagens da família de funções $f_\phi|_{[1,+\infty)}$, $\phi \in [0, 2\pi)$, e $f_\phi|_{(-\infty,-1]}$, $\phi \in [0, 2\pi)$, respectivamente, definidas por

$$f_\phi(r) \;=\; r - a_1(-1)^n\sqrt{r^2-1}\sin\phi, \qquad r \,\in\, (-\infty, -1] \cup [1, +\infty).$$

Para n par, como $f_{\pi/2} \leq f_\phi \leq f_{3\pi/2}$, qualquer que seja $\phi \in \mathbb{R}$, basta estudar os extremos das funções $f_{\pi/2}$ e $f_{3\pi/2}$ ($f_{\pi/2}$ não admite mínimo, nem $f_{3\pi/2}$ máximo, quando $a_1 \geq 1$). Se $a_1 < 1$, verifica-se que α_1' é o mínimo de $f_{\pi/2}|_{[1,+\infty)}$ e $f_{3\pi/2}|_{[1,+\infty)}$ não admite máximo, ou seja, $W_J^+(A_1) = [\alpha_1', +\infty)$. Se $a_1 = 1$, constata-se que $f_{\pi/2}(r) > 0$, para todo o $r \geq 1$, e $W_J^+(A_1) = (0, +\infty)$. Para n ímpar, invertem-se os papéis das funções $f_{\pi/2}$ e $f_{3\pi/2}$, mas tiram-se as mesmas conclusões. Mais se observa que $f_{\pi/2}(-r) = -f_{3\pi/2}(r)$, o que permite descrever $-W_J^-(A_1)$ a partir de $W_J^+(A_1)$. Em resumo,

i. $W_J(A_1) = \mathbb{R} \setminus]-\alpha_1', \alpha_1'[$, se $a_1 < 1$;

ii. $W_J(A_1) = \mathbb{R} \setminus \{0\}$, se $a_1 = 1$;

iii. $W_J(A_1) = \mathbb{R}$, se $a_1 > 1$ (pelo Teorema 1.2.4, pois A_1 é Hermítica-J e $\alpha_1' \in \mathbb{C} \setminus \mathbb{R}$).

b) Se $a_1 \neq 0$ e $\theta_1 \neq \pi/2 + n\pi$, $n \in \mathbb{Z}$, então caracterizam-se os elementos de $W_J(A_1)$ pela família de segmentos de recta $y \sin\theta_1 = (r - x)\cos\theta_1$, com $r \in \mathbb{R}$ tal que $r^2 - 1 \geq 0$, de extremos determinados por $y = \pm\, a_1\sqrt{r^2-1}\cos\theta_1$. Eliminando o parâmetro r destas duas equações, obtém-se

$$(x + \operatorname{tg}\theta_1\, y)^2 - \left(\frac{y}{a_1\cos\theta_1}\right)^2 \;=\; 1,$$

precisamente a equação da cónica $\mathrm{A}x^2 + 2\mathrm{B}xy + \mathrm{C}y^2 + \mathrm{F} = 0$ obtida em I, com A,B,C,F definidos por (1.26), para o caso particular $a_1 = b_1$. Assim, as conclusões para $W_J(A_1)$ advêm das obtidas em I, ou seja, $W_J(A_1)$ é uma hipérbole como a do subcaso I.a), atendendo a que $M_1, N_1 > 0$, quando se toma $a_1 = b_1$ em (1.29) e (1.30).

III. Seja $k = 0$. Para $a_0 = b_0 = 0$, é claro que $W_J(A_0) = \{0\}$. Se $a_0 = b_0 \neq 0$, então $W_J(A_0) = ie^{i\theta_0}\mathbb{R}$, a recta que passa pela origem e é perpendicular à recta definida pelos valores próprios de A_0. Se $a_0 \neq b_0$, observa-se que $\mathrm{e}^{-i\theta_0}(x + i\,y) \in W_J(A_0)$ se e só se

$$\frac{x^2}{(a_0+b_0)^2} + \frac{y^2}{(a_0-b_0)^2} \;=\; \frac{r^2-1}{4}, \qquad r \,\in\, (-\infty, -1] \cup [1, +\infty),$$

ou seja, $W_J(A_0) = \mathbb{C}$.

IV. Concluída a descrição de $W_J(A_k)$, $k = 0, 1$, podem tirar-se as devidas conclusões para $W_J(A)$. Ora, resulta de (1.24), pelas propriedades $\mathrm{W_H}1$ e $\mathrm{W_H}2$, que

$$W_J(A) \;=\; \alpha \,+\, \beta_k W_J(A_k),$$

com $k = 0$, se $a_{11} = a_{22}$, e $k = 1$, caso contrário. Constata-se, ainda, que $W_J(A)$ contém os valores próprios α_1 e α_2 de A, desde que $W_J(A_k)$ contenha os valores próprios α_1' e $-\alpha_1'$ de A_k, uma vez que

$$\alpha_1 \;=\; \alpha + \beta_k \alpha_1' \qquad \text{e} \qquad \alpha_2 \;=\; \alpha - \beta_k \alpha_1'. \tag{1.32}$$

Da relação (1.32) entre os valores próprios de A e A_k e a partir de (1.24), donde resulta $U^{[*]}A^{[*]}U = \bar{\alpha} I_2 + \bar{\beta}_k A_k^{[*]}$, concluem-se as igualdades

$$\mathcal{Q}_A \;=\; 2|\alpha|^2 \;+\; |\beta_k|^2 \mathcal{Q}_{A_k}, \tag{1.33}$$

onde $(\mathcal{Q}_A, \mathcal{Q}_{A_k})$ é dado pelas expressões $(|\alpha_1|^2 + |\alpha_2|^2, 2|\alpha_1'|^2)$ ou $(2\mathrm{Re}\,(\overline{\alpha}_1\alpha_2)\,, 2|\alpha_1'|^2)$ ou $(\mathrm{Tr}(A^{[*]}A), \mathrm{Tr}(A_k^{[*]}A_k))$. A partir de (1.33), mediante cálculos elementares, verifica-se que M, N definidos em (1.23) e M_1, N_1 em (1.31) estão relacionados da seguinte forma:

$$M \;=\; |\beta_1|^2 M_1 \qquad \text{e} \qquad N \;=\; |\beta_1|^2 N_1.$$

Assim, se $a_{11} = a_{22}$, a partir do estudo feito para $k = 0$ em III, conclui-se que $W_J(A)$ é o ponto α, se A é uma matriz escalar; a recta l que passa por α e que é perpendicular à recta definida por α_1 e α_2, se $|a_{12}| = |a_{21}| \neq 0$; e todo o plano complexo, caso contrário.

Agora, seja $a_{11} \neq a_{22}$ e utilize-se o estudo antes efectuado para $k = 1$. Quando $|a_{12}| \neq |a_{21}|$, com base no discutido em I, caso em que $M > 0$, tem-se:

A) Se $N > 0$, então $W_J(A)$ é limitado por uma hipérbole centrada em α, de focos α_1 e α_2, eixos transverso e não-transverso de comprimento $\sqrt{N}$ e $\sqrt{M}$, respectivamente;

B) Se $N = 0$, então $W_J(A)$ é todo o plano complexo, excepto a recta l;

C) Se $N < 0$, então $W_J(A)$ é todo o plano complexo.

Quando $|a_{12}| = |a_{21}| \neq 0$ e a diferença entre $\arg(a_{12}a_{21})$ e $2\arg(a_{11} - a_{22})$ é um ângulo distinto de π, então $M, N > 0$ e, por II.b), o conjunto $W_J(A)$ tem a fronteira hiperbólica descrita em A). Quando $|a_{12}| = |a_{21}|$ e $\arg(a_{12}a_{21})$ e $2\arg(a_{11} - a_{22})$ diferem entre si por um ângulo π, ou A é uma matriz diagonal não escalar, então $W_J(A)$ é um subconjunto da recta s, dada por $\alpha + \beta_1\mathbb{R}$, que passa por α e é claramente definida por $a_{11} = \alpha + \beta_1$ e $a_{22} = \alpha - \beta_1$. A partir da análise tecida em II.a), conclui-se, em particular, que:

D) Se $2|a_{12}| < |a_{11} - a_{22}|$, então $W_J(A)$ é a recta s, excepto o segmento aberto que une α_1 a α_2 (nota-se que α_1, α_2 são distintos e pertencem a s);

E) Se $2|a_{12}| = |a_{11} - a_{22}|$, então $W_J(A)$ é toda a recta s, excepto o ponto α (nota-se que $\alpha = \alpha_1 = \alpha_2$);

F) Se $2|a_{12}| > |a_{11} - a_{22}|$, então $W_J(A)$ é toda a recta s (nota-se que α_1, α_2 são distintos, definem uma recta perpendicular a s e, portanto, s coincide com l).

O caso descrito em E) ocorre, se $M = N = 0$. Nos restantes casos, em que $W_J(A)$ é um subconjunto de uma recta, pelo menos uma das constantes M ou N é zero, sendo a outra positiva. Perante a união de semi-rectas em D), temos $M = 0$ e $N > 0$. Quando a situação da recta em F) ocorre, temos $N = 0$ e $M > 0$. ■

A prova que apresentamos para o Teorema do Contradomínio Hiperbólico é algo extensa. O tratamento exaustivo que fazemos tem a virtude de permitir discernir, de forma consistente, a razão da ocorrência de cada conjunto degenerado.

Nota 1.2.2 Perante as hipóteses do Teorema 1.2.7 a), b) ii e d), a matriz $A \in M_2$ admite um valor próprio em $\sigma_J^+(A)$, digamos α_1. Da demonstração anterior, facilmente se deduz que o conjunto $W_J^+(A)$ é:

- a componente convexa de $W_J(A)$, limitada pelo ramo da hipérbole, que contém o valor próprio α_1, na hipótese a);
- o semi-plano aberto definido pela recta l e contendo α_1, na hipótese b) ii;
- a semi-recta fechada em $W_J(A)$ com α_1 por extremo, na hipótese d);
- a semi-recta aberta de extremo α, contendo o ponto a_{11}, na hipótese e) ii;
- coincidente com $W_J(A)$, nos restantes casos.

Como mencionado no ponto III da Nota 1.2.1, estamos, agora, em condições de completar a caracterização de $W_H(A)$, quaisquer que sejam $A \in M_2$ e $H \in H_2$ indefinida não-singular. Para tal, basta recordar a relação (1.14) entre o contradomínio numérico-H de A e o contradomínio numérico-J de $A_R = R^{-1}AR$, onde $R \in M_2$ é uma matriz não-singular, satisfazendo $R^*HR = J$. Observa-se que os valores próprios de A e A_R coincidem, que $\mathrm{Tr}(H^{-1}A^*HA) = \mathrm{Tr}(JA_R^*JA_R)$ e aplica-se o Teorema 1.2.7. Conclui-se, assim, que $W_H(A)$ é limitado por uma hipérbole, possivelmente degenerada, de focos nos valores próprios α_1 e α_2 de A, cujos eixos transverso e não-transverso têm comprimento $\sqrt{N_H}$ e $\sqrt{M_H}$, respectivamente, onde

$$M_H \;=\; |\alpha_1|^2 + |\alpha_2|^2 - \mathrm{Tr}\big(A^{[*]}A\big)\,, \qquad N_H \;=\; \mathrm{Tr}\big(A^{[*]}A\big) - 2\,\mathrm{Re}\,(\bar{\alpha}_1\alpha_2)\,,$$

e $A^{[*]}$ é a adjunta-H de A. Nos casos degenerados, $W_H(A)$ reduz-se a um ponto, a um subconjunto de uma recta, a todo o plano complexo, ou a todo o plano excepto uma recta. Os sinais das constantes M_H e N_H (e a relação entre as entradas diagonais ou não diagonais de A_R) ditam os diferentes casos degenerados.

Como corolário do Teorema 1.2.7, obtém-se a descrição de $W_J^+(A)$, para matrizes $A \in M_2$ Hermíticas-J.

Corolário 1.2.4 *Se* $J = \operatorname{diag}(1, -1)$ *e* $A = [a_{ij}] \in M_2$ *é uma matriz Hermítica-J de valores próprios* α_1 *e* α_2*, então* $W_J^+(A)$ *é:*

i) $\mathbb{R}$, *se* $\alpha_1, \alpha_2 \in \mathbb{C} \setminus \mathbb{R}$*;*

ii) $[\alpha_1, +\infty)$, *se* $\alpha_1 > \alpha_2$ *e* $a_{11} > a_{22}$*;*

iii) $(-\infty, \alpha_1]$, *se* $\alpha_1 < \alpha_2$ *e* $a_{11} < a_{22}$*;*

iv) $\{\alpha_1\}$, *se* $\alpha_1 = \alpha_2$ *e* $a_{11} = a_{22}$*;*

v) $(\alpha_1, +\infty)$, *se* $\alpha_1 = \alpha_2$ *e* $a_{11} > a_{22}$*;*

vi) $(-\infty, \alpha_1)$, *se* $\alpha_1 = \alpha_2$ *e* $a_{11} < a_{22}$*;*

com $\alpha_1, \alpha_2 \in \mathbb{R}$ *nos casos* ii) - vi).

Demonstração: Repetindo o racíocinio inicial da demonstração do Teorema 1.2.7, o Lema 1.2.3 garante que existe uma matriz pseudo-unitária $U \in U_{1,1}$, satisfazendo

$$U^{[*]}AU \;=\; \alpha I_2 \;+\; \beta_k A_k, \tag{1.34}$$

onde

$$\alpha \;=\; \frac{1}{2}\mathrm{Tr}(A), \qquad \beta_k \;=\; \frac{1}{2}(a_{11} - a_{22})k + 1 - k \qquad \text{e} \qquad A_k \;=\; \begin{bmatrix} k & a_k \mathrm{e}^{i\,\theta_k} \\ b_k \mathrm{e}^{i\,\theta_k} & -k \end{bmatrix},$$

em que

$$a_k \;=\; |a_{12}||\beta_k|^{-1}, \qquad b_k \;=\; |a_{21}||\beta_k|^{-1}, \qquad \theta_k \;=\; \frac{1}{2}\arg(a_{12}a_{21}) - \arg\beta_k,$$

com $k = 0$, se $a_{11} = a_{22}$, e $k = 1$, caso contrário. Claramente, A é Hermítica-J se e só se A_k é Hermítica-J se e só se $a_k = b_k$ e ($a_k = 0$ ou $\theta_k = \pi/2 + n\pi$, $n \in \mathbb{Z}$). Nestas condições, $\alpha, \beta_k \in \mathbb{R}$ e os valores próprios de A_k são dados por $\alpha_1' = \sqrt{k^2 - a_k^2}$ e $\alpha_2' = -\alpha_1'$. Além disso, como os valores próprios de A e de A_k satisfazem a relação $\alpha_j = \alpha + \beta_k \alpha_j'$, $j = 1, 2$, a partir do estudo pormenorizado efectuado nos casos II.a) e III da demonstração do Teorema 1.2.7, podemos clarificar as diversas situações que podem surgir:

i) Obviamente, $\alpha_j \notin \mathbb{R}$ se e só se $\alpha_1' \in i\,(\mathbb{R} \setminus \{0\})$, o que ocorre quando $k = 0$ e $a_0 \neq 0$, ou quando $k = 1$ e $a_1 < 1$. Em ambos os casos, resulta da demonstração do Teorema 1.2.7 que $W_J^+(A_k) = \mathbb{R}$, $k = 0, 1$.

ii) - iii) Ora, $\alpha_j \in \mathbb{R}$ se e só se $\alpha_1' \in \mathbb{R}$. Além disso, $\alpha_1 \neq \alpha_2$ se e só se $\alpha_1' \neq 0$. Tal ocorre apenas quando $k = 1$ e $a_1 > 1$, verificando-se que $W_J^+(A_1) = [\sqrt{1 - a_1^2}, +\infty)$.

iv) A matriz A é escalar se e só se $k = 0$ e $A_0 = 0$, tendo-se $W_J^+(A_0) = \{0\}$.

v) - vi) A matriz A é não escalar e $\alpha_1 = \alpha_2$ se e só se $k = 1$ e $\alpha_1' = 0$, isto é, $a_1 = 1$, caso em que $W_J^+(A_1) = (0, +\infty)$.

As propriedades $\mathrm{W_H}1$ e $\mathrm{W_H}2$' aplicadas a (1.34) conduzem a $W_J^+(A) = \alpha + \beta_k W_J^+(A_k)$. Facilmente se caracteriza $W_J^+(A)$ a partir do já estudado conjunto $W_J^+(A_k)$, que vem agora afectado do sinal de β_1, quando $k = 1$, permitindo distinguir entre ii) e iii), e entre v) e vi). ■

A partir do Corolário 1.2.4, pode descrever-se, com facilidade, o contradomínio numérico-J de matrizes $A \in M_2$ essencialmente Hermíticas-J.

Exemplo 1.2.2 Sejam $J = \operatorname{diag}(1, -1)$ e a seguinte matriz essencialmente Hermítica-J

$$A \;=\; (1+i)\begin{bmatrix} 1 & a \\ -a & -1 \end{bmatrix}, \qquad a > 0.$$

Designa-se por $l_{\pi/4}$ a bissectriz dos quadrantes ímpares do plano complexo. Se $a > 1$, então $W_J(A)$ é a recta $l_{\pi/4}$. Se $a = 1$, então $W_J(A)$ é a recta $l_{\pi/4}$, com excepção da origem. Se $a < 1$, apenas os pontos $\pm\, \mathrm{e}^{i\pi/4}\sqrt{1-a^2}$ pertencem à fronteira de $W_J(A)$. Em particular, quando $a = \sqrt{2}/2$, tem-se que $W_J(A)$ é a união disjunta das semi-rectas fechadas, colineares a $l_{\pi/4}$, de extremos $\frac{1}{2}(1-i)$ e $\frac{1}{2}(1+i)$.

1.2.5 Contradomínios Hiperbólicos de Matrizes por Blocos

Considera-se, para $0 < r < n$, a matriz por blocos

$$A \;=\; \begin{bmatrix} aI_r & X \\ Y & bI_{n-r} \end{bmatrix}, \qquad a, b \in \mathbb{C}, \quad X, Y^* \in M_{r\times(n-r)}, \tag{1.35}$$

em que XY e YX são matrizes normais, e $p = \min(r, n-r)$. Ethan S. Brown e Ilya M. Spitkovsky [36] provaram que $W(A)$ é o invólucro convexo de p elipses, algumas eventualmente coincidentes e/ou degeneradas, todas centradas em $(a+b)/2$. Para A do tipo (1.35) e $J = I_r \oplus -I_{n-r}$, provaremos que o conjunto $W_J(A)$ é, sob determinadas condições, o invólucro pseudo-convexo de p hipérboles não degeneradas, todas centradas em $\frac{1}{2}(a+b)$.

Lema 1.2.4 *Seja A uma matriz por blocos do tipo* (1.35), $0 < r < n$, *com XY e YX matrizes normais, e $p = \min(r, n-r)$. Sejam $\sigma_1, \ldots, \sigma_p$ e $\delta_1, \ldots, \delta_p$ os valores singulares de X e Y, respectivamente. Então existe uma matriz pseudo-unitária $U \in U_{r,n-r}$, tal que*

$$U^{-1}AU \;=\; \begin{bmatrix} aI_r & \Sigma \\ \Delta & bI_{n-r} \end{bmatrix}, \tag{1.36}$$

onde Σ e Δ são matrizes de entradas não diagonais nulas e de entradas da diagonal principal dadas por $\sigma_1 \mathrm{e}^{i\phi_1}, \ldots, \sigma_p\, \mathrm{e}^{i\phi_p}$ e $\delta_1 \mathrm{e}^{i\phi_1}, \ldots, \delta_p\, \mathrm{e}^{i\phi_p}$, respectivamente, para determinados $\phi_1, \ldots, \phi_p \in [0, 2\pi)$.

Demonstração: Pelo Teorema da Decomposição dos Valores Singulares [80], existem matrizes unitárias $U_1 \in M_r$ e $U_2 \in M_{n-r}$, tais que $U_1^* X U_2$ tem as entradas da diagonal principal dadas pelos valores singulares $\sigma_1, \ldots, \sigma_p$ de X e as restantes entradas são nulas. Sendo XY e YX ambas matrizes normais, existem $\phi_1, \ldots, \phi_p \in [0, 2\pi)$, tais que os elementos da diagonal principal de $U_1^* Y^* U_2$ são dados por $\delta_1 \mathrm{e}^{2i\phi_1}, \ldots, \delta_p \mathrm{e}^{2i\phi_p}$ e os restantes elementos são zero [80, p.426]. Tomando

$$D_1 \; = \; \operatorname{diag}\left(\mathrm{e}^{i\eta_1}, \ldots, \mathrm{e}^{i\eta_p}\right) \oplus I_{r-p}, \qquad D_2 \; = \; \operatorname{diag}\left(\mathrm{e}^{i\mu_1}, \ldots, \mathrm{e}^{i\mu_p}\right) \oplus I_{n-r-p},$$

em que $\mu_k - \eta_k = \phi_k$, $k = 1, \ldots, p$, verifica-se, com facilidade, que a matriz por blocos $U = (U_1 D_1) \oplus (U_2 D_2)$ satisfaz simultaneamente $U^* J U = J$, $J = I_r \oplus -I_{n-r}$, e (1.36). ■

Antes da apresentação do resultado principal desta secção, relembra-se que se $T \in M_n$ é uma matriz por blocos do tipo

$$T \; = \; \begin{bmatrix} X & Y \\ Z & W \end{bmatrix}, \qquad X \in M_r, \quad W \in M_{n-r},$$

em que X é não-singular, então $\det(T) = \det(X)\,\det(W - ZX^{-1}Y)$ [89]. Em particular, se $n = 2r$ e X, Y, Z, W são matrizes que comutam entre si, então $\det(T) = \det(XW - ZY)$.

Teorema 1.2.8 *Seja $J = I_r \oplus -I_{n-r}$, $0 < r < n$, seja A uma matriz por blocos do tipo (1.35), com XY e YX matrizes normais, e $p = \min(r, n-r)$. Consideram-se*

$$\beta_{j\pm} \; = \; \frac{1}{2}(a+b) \, \pm \, \frac{1}{2}\sqrt{(a-b)^2 + 4\,\sigma_j \delta_j \mathrm{e}^{2i\phi_j}}, \qquad j = 1, \ldots, p, \tag{1.37}$$

em que $\sigma_1, \ldots, \sigma_p$ e $\delta_1, \ldots, \delta_p$ são os valores singulares de X e Y, respectivamente, e $\phi_1, \ldots, \phi_p \in [0, 2\pi)$ são determinados por X e Y como no Lema 1.2.4. Se

$$2\,\mathrm{Re}\left(\bar{\beta}_{j+}\beta_{j-}\right) \; < \; |a|^2 + |b|^2 - \sigma_j^2 - \delta_j^2 \; < \; |\beta_{j+}|^2 + |\beta_{j-}|^2, \qquad j = 1, \ldots, p, \tag{1.38}$$

então a curva geradora de fronteira de $W_J(A)$ é constituída por p hipérboles não degeneradas (algumas possivelmente coincidentes), todas centradas em $\frac{1}{2}(a+b)$, de focos β_{j+} e β_{j-}, eixo não-transverso de comprimento

$$\sqrt{|\beta_{j+}|^2 + |\beta_{j-}|^2 - |a|^2 - |b|^2 + \sigma_j^2 + \delta_j^2}, \qquad j = 1, \ldots, p, \tag{1.39}$$

e eventualmente um ponto, a se $n < 2r$, e b se $n > 2r$. O conjunto $W_J(A)$ é o invólucro pseudo-convexo destas p hipérboles.

Demonstração: Pelo Lema 1.2.4, atendendo à propriedade $\mathrm{W_H}1$, pode centrar-se a atenção no estudo de $W_J(B)$, em que B é a matriz por blocos em (1.36). O polinómio característico de

$$\mathrm{Re}^J\big(\mathrm{e}^{i\theta}B\big) \;=\; \frac{1}{2}\begin{bmatrix} 2\,\mathrm{Re}\big(a\mathrm{e}^{i\theta}\big)I_r & \mathrm{e}^{i\theta}\Sigma - \mathrm{e}^{-i\theta}\Delta^* \\ \mathrm{e}^{i\theta}\Delta - \mathrm{e}^{-i\theta}\Sigma^* & 2\,\mathrm{Re}\big(b\mathrm{e}^{i\theta}\big)I_{n-r} \end{bmatrix},$$

é dado por

$$P_n(t) \;=\; (t-c_{a\theta})^{r-p}(t-c_{b\theta})^{n-r-p}\,\det\left((t-c_{a\theta})(t-c_{b\theta})I_p - D_\theta\right),$$

em que $c_{a\theta} = |a|\cos(\theta+\arg a)$, $c_{b\theta} = |b|\cos(\theta+\arg b)$ e D_θ é a matriz diagonal $p\times p$, cuja j-ésima entrada diagonal é

$$d_j^\theta \;=\; \frac{1}{2}\sigma_j\delta_j\cos(2\theta+2\phi_j) - \frac{1}{4}(\sigma_j^2+\delta_j^2), \qquad j=1,\ldots,p.$$

As raízes do polinómio $P_n(t)$ satisfazem

$$(t-c_{a\theta})^{r-p}(t-c_{b\theta})^{n-r-p}\prod_{j=1}^{p}\left(t^2-(c_{a\theta}+c_{b\theta})\,t + c_{a\theta}c_{b\theta} - d_j^\theta\right) \;=\; 0,$$

e os valores próprios de $\mathrm{Re}^J\big(\mathrm{e}^{i\theta}B\big)$ são

$$\lambda_{j\pm}^\theta \;=\; \frac{1}{2}\left(c_{a\theta}+c_{b\theta}\right) \pm \frac{1}{2}\sqrt{\left(c_{a\theta}-c_{b\theta}\right)^2 + 4d_j^\theta}, \qquad j=1,\ldots,p, \tag{1.40}$$

e $c_{a\theta}$ $(c_{b\theta})$, se $n<2r$ $(n>2r)$. Obviamente σ_j, δ_j não são simultaneamente zero, caso contrário a segunda desigualdade em (1.38) não se verificaria. O vector $u_{j\pm}^\theta \in \mathbb{C}^n$, de componentes j e $r+j$ iguais a $2c_{a\theta} - 2\lambda_{j\pm}^\theta$ e $\sigma_j\mathrm{e}^{-i(\theta+\phi_j)} + \delta_j\mathrm{e}^{i(\theta+\phi_j)}$, respectivamente, todas as restantes sendo zero, é um vector próprio de $\mathrm{Re}^J\big(\mathrm{e}^{i\theta}B\big)$ associado ao valor próprio $\lambda_{j\pm}^\theta$, $j=1,\ldots,p$. Se $n<2r$, o vector e_l da base canónica de $\mathbb{C}^n$ é um vector próprio de $\mathrm{Re}^J\big(\mathrm{e}^{i\theta}B\big)$ associado ao valor próprio $c_{a\theta}$, que satisfaz $\langle Be_l, e_l\rangle_J = a$, $l=n-r+1,\ldots,r$; se $n>2r$, então e_l é um vector próprio de $\mathrm{Re}^J\big(\mathrm{e}^{i\theta}B\big)$ associado ao valor próprio $c_{b\theta}$, satisfazendo $\langle Be_l, e_l\rangle_J = -b$, $l=2r+1,\ldots,n$. Como (1.38) se verifica, existem direcções determinadas pelos ângulos θ, para as quais todos os valores próprios de $\mathrm{Re}^J\big(\mathrm{e}^{i\theta}B\big)$ são números reais. Assim, a partir do Teorema 1.2.3, conclui-se que a curva geradora de fronteira de $W_J(B)$ coincide com as curvas geradoras de fronteira de $W_{J_2}(B_j)$, quando $J_2 = \mathrm{diag}\,(1,-1)$ e

$$B_j \;=\; \begin{bmatrix} a & \sigma_j\mathrm{e}^{i\phi_j} \\ \delta_j\mathrm{e}^{i\phi_j} & b \end{bmatrix}, \qquad j=1,\ldots,p, \tag{1.41}$$

e possivelmente um ponto, a se $n<2r$, e b se $n>2r$. Os valores próprios de B_j são $\beta_{j\pm}$ definidos em (1.37) e vale $\mathrm{Tr}\,(B_j^{[*]}B_j) = |a|^2+|b|^2-\sigma_j^2-\delta_j^2$. Pelo Teorema 1.2.7 a), conclui-se que $W_{J_2}(B_j)$ é limitado por uma hipérbole, de focos em β_{j+}, β_{j-} e eixo não--transverso de comprimento (1.39). Se $n = 2r$, a curva geradora de fronteira de $W_J(B)$ é dada precisamente por estas p hipérboles, algumas das quais podem coincidir. Além disso, é claro que $a \in W^+_{J_2}(B_j)$ e $b \in -W^-_{J_2}(B_j)$, $j = 1, \ldots, p$. Consequentemente, mesmo quando as matrizes X, Y são rectangulares, isto é, no caso $n \neq 2r$, o conjunto $W_J(B)$ é o invólucro pseudo-convexo das p hipérboles anteriormente descritas. ■

Exemplo 1.2.3 Sejam $J = I_2 \oplus -I_2$ e

$$A = \begin{bmatrix} 1 & 0 & \frac{1}{2} & 0 \\ 0 & 1 & 0 & \frac{3}{2} \\ 1 & 0 & -1 & 0 \\ 0 & i & 0 & -1 \end{bmatrix}. \tag{1.42}$$

Ora, A é uma matriz por blocos do tipo (1.35), com $a = 1$, $b = -1$, tal que $XY = YX$ é uma matriz diagonal. Pelo Teorema 1.2.8, $W_J(A)$ é o invólucro pseudo-convexo de duas hipérboles apresentadas na Figura 1.1.

Figura 1.1: Curva geradora de fronteira de $W_J(A)$, para A e J no Exemplo 1.2.3.

Considera-se, agora, $Y^* = kX$, $k \in \mathbb{C}$, na matriz A definida por blocos em (1.35). Neste caso especial, A admite um contradomínio numérico clássico elíptico [36]. No corolário seguinte, obtém-se o paralelo hiperbólico para o contradomínio numérico-J.

Corolário 1.2.5 *Seja $J = I_r \oplus -I_{n-r}$, $0 < r < n$, seja A uma matriz por blocos do tipo* (1.35), *com $Y^* = kX$, $k \in \mathbb{C}$, e $p = \min(r, n-r)$. Consideram-se*

$$\beta_{j\pm} = \frac{1}{2}(a+b) \pm \frac{1}{2}\sqrt{(a-b)^2 - 4k\sigma_j^2}, \qquad j = 1, \ldots, p,$$

em que $\sigma_1, \ldots, \sigma_p$ *são os valores singulares de* X. *Se*

$$2\,\mathrm{Re}\left(\bar{\beta}_{j+}\beta_{j-}\right) \;<\; |a|^2 + |b|^2 - \sigma_j^2(1+|k|^2) \;<\; |\beta_{j+}|^2 + |\beta_{j-}|^2, \qquad j = 1, \ldots, p,$$

então $W_J(A)$ *é limitado por uma hipérbole, de focos* β_{1+} *e* β_{1-}, *cujo eixo não-transverso tem comprimento*

$$\sqrt{|\beta_{1+}|^2 + |\beta_{1-}|^2 - |a|^2 - |b|^2 + \sigma_1^2(1+|k|^2)}.$$

Demonstração: Da prova do Teorema 1.2.8, conclui-se que $W_J^+(A)$ é o invólucro convexo de $W_{J_2}^+(B_j)$, onde $J_2 = \mathrm{diag}\,(1,-1)$ e B_j são as matrizes 2×2 em (1.41), $j = 1, \ldots, p$. Como $Y^* = kX$, então $\delta_j = |k|\,\sigma_j$ e $2\phi_j = \arg k + \pi$, $j = 1, \ldots, p$. Provaremos que

$$W_{J_2}^+(B_j) \;\subseteq\; W_{J_2}^+(B_1), \qquad j = 2, \ldots, p. \tag{1.43}$$

Recordando a definição de $W_{J_2}^+(B_j)$, tem-se

$$W_{J_2}^+(B_j) \;=\; \left\{a|x_1|^2 - b|x_2|^2 + \sigma_j e^{i\phi_j}(\bar{x}_1 x_2 + |k|\bar{x}_2 x_1) :\; |x_1|^2 - |x_2|^2 = 1\right\}. \tag{1.44}$$

Tomando $r = |x_1|^2$ e $\gamma = \arg x_2 - \arg x_1$ em (1.44), e denotando por $\mathcal{E}$ a curva

$$\{(1+|k|)\cos\gamma + i(1-|k|)\sin\gamma :\; \gamma \in [0, 2\pi)\},$$

podemos escrever

$$W_{J_2}^+(B_j) \;=\; \bigcup_{r\geq 1}\left((a-b)r + b + \sigma_j e^{i\phi_j}\sqrt{r(r-1)}\,\mathcal{E}\right).$$

Qualquer $z \in W_{J_2}^+(B_j)$ está sobre a curva $\Gamma_j = (a-b)r + b + \sigma_j e^{i\phi_j}\sqrt{r(r-1)}\,\mathcal{E}$, para uma determinada escolha de r. Consoante o valor de k, a curva $\mathcal{E}$ é a fronteira de uma elipse centrada na origem, ou um segmento de recta. Como $\sigma_1 \geq \sigma_j$, então Γ_j está no domínio limitado por Γ_1. Ora $\Gamma_1 \subseteq W_{J_2}^+(B_1)$ e este último conjunto é convexo, pelo que $z \in \Gamma_j \subseteq W_{J_2}^+(B_1)$ e verifica-se (1.43). Assim, $W_J^+(A) = W_{J_2}^+(B_1)$. Resultado análogo vale para $W_J^-(A)$ e o corolário é consequência do Teorema 1.2.8. ■

Exemplo 1.2.4 Sejam $J = I_2 \oplus -I_2$ e

$$A \;=\; \begin{bmatrix} 1 & 0 & \frac{1}{2} & 0 \\ 0 & 1 & 0 & \frac{3}{2} \\ -\frac{i}{2} & 0 & -1 & 0 \\ 0 & -\frac{3i}{2} & 0 & -1 \end{bmatrix}. \tag{1.45}$$

Ora, A é uma matriz por blocos do tipo (1.35), com $a = 1$ e $b = -1$, $Y^* = iX$. Pelo Corolário 1.2.5, $W_J(A)$ é uma hipérbole e seu interior. A Figura 1.2 apresenta as duas hipérboles que formam a curva geradora de fronteira de $W_J(A)$.

Figura 1.2: Curva geradora de fronteira de $W_J(A)$, para A e J no Exemplo 1.2.4.

Chama-se *matriz quadrática* a uma matriz $A \in M_n$ que satisfaça uma equação quadrática do tipo $A^2+\beta A+\gamma I_n = 0$, com $\beta, \gamma \in \mathbb{C}$. Estas matrizes têm polinómio característico de grau dois. Se $A \in M_n$ é uma matriz quadrática de valores próprios a e b, então A é unitariamente semelhante a uma matriz do tipo

$$\begin{bmatrix} aI_r & X \\ 0 & bI_{n-r} \end{bmatrix}, \tag{1.46}$$

em que X é uma matriz semi-definida positiva. Shu-Hsien Tso e Pei Yuan Wu [128] mostraram que o contradomínio numérico clássico de matrizes quadráticas é um disco elíptico. Agora, ao tomar-se $k = 0$ no Corolário 1.2.5, obtém-se o resultado seguinte, especificando o contradomínio numérico-J desta classe especial de matrizes.

Corolário 1.2.6 *Seja $J = I_r \oplus -I_{n-r}$, $0 < r < n$, e seja A uma matriz quadrática unitariamente semelhante a uma matriz por blocos do tipo* (1.46). *Se $0 < \sigma_1 < |a - b|$, em que σ_1 é o maior valor próprio de X, então $W_J(A)$ é limitado por uma hipérbole de focos em a e b, e eixo não-transverso de comprimento σ_1.*

1.3 Contradomínio Tracial-C, H

Introduzimos, de seguida, uma das generalizações mais famosas do contradomínio numérico clássico, devida a Moshe Goldberg e E. G. Straus [67]. Dada $C \in M_n$, o *contradomínio numérico-C* de $A \in M_n$ é o subconjunto do plano complexo que se denota e define por

$$W_C(A) = \{\mathrm{Tr}\,(CU^*AU) : U \in M_n,\ U^*U = I_n\}.$$

Se $C = \mathrm{diag}\,(1, 0, \ldots, 0)$, então $W_C(A) = W(A)$.

Define-se a órbita unitária da matriz $A \in M_n$ por

$$\mathcal{O}(A) \;=\; \{U^*AU :\; U \in M_n,\; U^*U = I_n\}.$$

Pode encarar-se $\mathcal{O}(A)$ como o conjunto de todas as representações matriciais de um determinado operador linear L em $\mathbb{C}^n$, em relação a diferentes bases ortonormadas. Para ilustrar o seu papel, observa-se que o operador linear L é normal (auto-adjunto) se e só se $\mathcal{O}(A)$ contém uma matriz diagonal (diagonal real), ou que L é um múltiplo escalar do operador identidade se e só se $\mathcal{O}(A)$ é um conjunto singular.

Neste ponto, recorda-se que o produto interno usual da álgebra M_n é da forma

$$\langle X, Y\rangle \;=\; \mathrm{Tr}\,(Y^*X)\,, \qquad X,\, Y \,\in\, M_n,$$

Pelo Teorema da Representação de Riesz [45, 3.4], dada uma funcional linear F em M_n, existe uma matriz $C \in M_n$, tal que $F(X) = \mathrm{Tr}\,(CX)$, qualquer que seja $X \in M_n$. Assim, $W_C(A)$ é justamente a imagem da órbita unitária de $A \in M_n$ por uma funcional linear em M_n.

Ao substituir C por matrizes com estruturas especiais, $W_C(A)$ reduz-se a outros contradomínios numéricos mais simples. Apresentamos alguns exemplos.

Se $C = I_k \oplus 0_{n-k}$, então $W_C(A)$ diz-se o *contradomínio numérico-k* de $A \in M_n$ e denota-se por $W_k(A)$, $k = 1, \ldots, n$. Obviamente, $W_1(A) = W(A)$ e $W_n(A) = \{\mathrm{Tr}\,(A)\}$. O conjunto $W_k(A)$ é convexo. A primeira solução para este problema, proposto por Paul R. Halmos, deve-se a C. A. Berger [30] (consultar [73] para uma versão simplificada).

Se $C = \mathrm{diag}\,(c_1, \ldots, c_n)$, então $W_C(A)$ diz-se o *contradomínio numérico-c* de $A \in M_n$, simplesmente denotado por $W_c(A)$, em que $c = (c_1, \ldots, c_n) \in \mathbb{C}^n$. Este conceito foi introduzido, em 1975, por Westwick [131] que provou a convexidade de $W_c(A)$, $c \in \mathbb{R}^n$, usando teoria de Morse. Desde então, ao contrário do Teorema de Toeplitz-Hausdorff, não surgiram alternativas de demonstração para o Teorema de Westwick, com excepção de uma prova puramente algébrica de Yiu-Tung Poon [116].

Verifica-se que o conjunto $W_C(A)$ é unitariamente invariante por transformações de semelhança unitária de A ou de C. Assim sendo, se $C \in M_n$ é normal e os seus valores próprios são as componentes do vector $c \in \mathbb{C}^n$, então $W_C(A) = W_c(A)$. Em particular, se $C \in M_n$ é essencialmente Hermítica ou, equivalentemente, normal com valores próprios colineares, então resulta do Teorema de Westwick que $W_C(A)$ é convexo, para $A \in M_n$ arbitrária. Em geral, $W_C(A)$ não é convexo, mesmo para $A, C \in M_3$ normais [131].

Não obstante, Wai-Shun Cheung e Nam-Kiu Tsing [38] mostraram que $W_C(A)$ é um conjunto estrelado relativamente ao ponto $\frac{1}{n}\,\mathrm{Tr}\,(A)\,\mathrm{Tr}\,(C)$. Um conjunto S diz-se *estrelado* relativamente a um ponto x se $\alpha x + (1-\alpha)y \in S$, para todo o $y \in S$ e $0 \le \alpha \le 1$.

Alvo de intensa investigação recente e com muitas propriedades interessantes (ver [92] para uma síntese dos principais resultados conhecidos à data), o conjunto $W_C(A)$ motiva a nossa próxima definição.

Sejam $C \in M_n$ e $H \in H_n$ uma matriz não-singular. O *contradomínio tracial-C, H* de $A \in M_n$ denota-se e define-se por

$$W_C^H(A) \;=\; \left\{\mathrm{Tr}\left(CU^{[*]}AU\right):\; U \in M_n,\; U^*HU = H\right\}. \tag{1.47}$$

Este conjunto reduz-se ao contradomínio numérico-C, quando $H = \pm I_n$.

As seguintes propriedades básicas do contradomínio tracial-C, H deduzem-se facilmente da definição:

$\mathbf{W_C^H1}$. $W^H_{V^{[*]}CV}(U^{[*]}AU) = W_C^H(A)$, para quaisquer matrizes U e V unitárias-H;

$\mathbf{W_C^H2}$. $W_C^H(\alpha I_n + \beta A) = \alpha \mathrm{Tr}(C) + \beta W_C^H(A)$, para quaisquer escalares $\alpha, \beta \in \mathbb{C}$;

$\mathbf{W_C^H3}$. $W^H_{C^{[*]}}(A^{[*]}) = \{\bar{z} :\; z \in W_C^H(A)\}$;

$\mathbf{W_C^H4}$. $W_C^H(A) = W_A^H(C)$, isto é, são simétricos os papéis das matrizes A e C.

Sem perda de generalidade, pode considerar-se na definição de $W_C^H(A)$, no lugar da matriz H, a matriz $J = I_r \oplus -I_{n-r}$, onde r é o número de valores próprios positivos de H, contando multiplicidades. Efectivamente, utilizando a *lei de inércia de Sylvester* [80], facilmente se verifica que

$$W_C^H(A) \;=\; W^J_{C_R}(A_R), \tag{1.48}$$

em que R é uma matriz não-singular tal que $R^*HR = J$ é a matriz de inércia de H, $A_R = R^{-1}AR$ e $C_R = R^{-1}CR$.

Pelo facto do grupo pseudo-unitário $U_{r,n-r}$ ser conexo [114] e de $W_C^J(A)$ ser o contradomínio da aplicacão contínua de $U_{r,n-r}$ em $\mathbb{C}$ definida por $U \mapsto \mathrm{Tr}\left(CU^{[*]}AU\right)$, conclui-se que $W_C^J(A)$ é um conjunto conexo, quaisquer que sejam $A, C \in M_n$. Sendo o grupo unitário $U_{n,0}$ compacto, análogo raciocínio garante a compacidade do conjunto $W_C(A)$, quaisquer que sejam $A, C \in M_n$.

Seja $\sigma = \sigma_1\sigma_2$ uma permutação em S_n, o grupo simétrico de grau n, tal que σ_1 e σ_2 são permutações de $1, \ldots, r$ e $r+1, \ldots, n$, respectivamente. Se α_i e c_i, $i = 1, \ldots, n$, são os valores próprios de A e C, respectivamente, os pontos

$$z_\sigma \;=\; \sum_{i=1}^{r} c_i \alpha_{\sigma_1(i)} + \sum_{i=r+1}^{n} c_i \alpha_{\sigma_2(i)}, \tag{1.49}$$

dizem-se os *pontos-σ* de $W_C^J(A)$. Quando $r = n$, os pontos z_σ, $\sigma \in S_n$, pertencem a $W_C(A)$ e desempenham um papel fundamental nesta teoria.

Uma matriz U de colunas $x_1, \dots, x_n \in \mathbb{C}^n$ pertence ao grupo pseudo-unitário $U_{r,n-r}$ se e só se o conjunto de vectores $x_1, \dots, x_n$ forma uma base *ortonormada-J*, isto é,

$$\langle x_k, x_l \rangle_J = \; j_k \, \delta_{kl}, \qquad k, l = 1, \dots, n, \tag{1.50}$$

onde j_k denota a k-ésima entrada diagonal da matriz $J = I_r \oplus -I_{n-r}$, $0 \le r \le n$.

Se $C = \operatorname{diag}(c_1, \dots, c_n) \in M_n$, então decorre directamente da definição do contradomínio tracial-C, J que

$$W_C^J(A) \;=\; \left\{ \sum_{k=1}^{n} j_k \, c_k \, \langle A x_k, x_k \rangle_J : \; x_1, \dots, x_n \in \mathbb{C}^n \text{ satisfazendo (1.50)} \right\}. \tag{1.51}$$

Podemos, sem perda de generalidade, fixar uma ordem para os elementos principais da matriz diagonal C, uma vez que a reordenação das primeiras r e das últimas $n-r$ entradas principais não afecta a forma de $W_C^J(A)$. Efectivamente, se $C_\sigma = P_\sigma^T C P_\sigma$, onde P_σ é a matriz de permutação associada a $\sigma = \sigma_1 \sigma_2 \in S_n$, com σ_1 e σ_2 permutações de $1, \dots, r$ e $r+1, \dots, n$, respectivamente, pela propriedade $\mathrm{W}_\mathrm{C}^\mathrm{H}1$, vale a igualdade $W_C^J(A) = W_{C_\sigma}^J(A)$.

Por exemplo, se $c = (c_1, \dots, c_n) \in \mathbb{R}^n$, como $W_c(A)$ não depende da ordenação das componentes de c, podemos supor a seguinte ordem não-crescente: $c_1 \ge \cdots \ge c_n$.

Em particular, se $C = E_{kk}$, onde $\{E_{11}, \dots, E_{nn}\}$ denota a base canónica de M_n, então $W_C^J(A)$ reduz-se a $\pm W_J^{\pm}(A)$, consoante $j_k = \pm 1$.

1.3.1 Formas Especiais do Contradomínio Tracial-C, J

Uma técnica muito útil na teoria de contradomínios numéricos e suas generalizações é a redução de certos problemas ao caso bidimensional. O primeiro e mais famoso problema a usufruir desta manipulação engenhosa foi a convexidade do contradomínio numérico clássico, ou seja, o Teorema de Toeplitz-Hausdorff.

O Caso 2×2 para $W_C^J(A)$ e C diagonal

Moshe Goldberg e E. G. Straus [67] caracterizaram o contradomínio numérico-C de matrizes 2×2, quando C é diagonal. Este Teorema do Contradomínio Elíptico relativo a $W_C(A)$ mostra que o conjunto é convexo, neste caso especial.

Teorema 1.3.1 *Se* $C = \operatorname{diag}(c_1, c_2) \in M_2$ *e* $A \in M_2$ *tem valores próprios* α_1 *e* α_2, *então* $W_C(A)$ *é um disco elíptico (possivelmente degenerado), cujos focos são os pontos-*σ, $z_{\text{id}} = c_1\alpha_1 + c_2\alpha_2$ *e* $z_{(12)} = c_1\alpha_2 + c_2\alpha_1$, *e cujos eixos maior e menor têm comprimento*

$$|c_1 - c_2|\sqrt{\operatorname{Tr}(A^*A) - 2\operatorname{Re}(\bar{\alpha}_1\alpha_2)} \qquad e \qquad |c_1 - c_2|\sqrt{\operatorname{Tr}(A^*A) - |\alpha_1|^2 - |\alpha_2|^2},$$

respectivamente.

Demonstração [67]**:** Como $C = c_2\, I_2 + (c_1 - c_2)\, E_{11}$ e $W_{E_{11}}(A) = W(A)$, atendendo às propriedades $\mathrm{W}_{\mathrm{C}}^{\mathrm{H}}2$ e $\mathrm{W}_{\mathrm{C}}^{\mathrm{H}}4$ (com $H = I_n$), tem-se

$$W_C(A) \;=\; c_2\operatorname{Tr}(A) + (c_1 - c_2)W(A). \tag{1.52}$$

Se C é uma matriz escalar, é óbvio que $W_C(A) = \{\operatorname{Tr}(A)\operatorname{Tr}(C)/2\}$. Se C é não-escalar, o resultado é imediato a partir de (1.52) e do Teorema 1.1.1. ■

Prossegue-se com a caracterização do contradomínio tracial-C, J para matrizes arbitrárias 2×2, quando C é diagonal e $J = \operatorname{diag}(1, -1)$, a qual decorre de modo simples e natural do Teorema do Contradomínio Hiperbólico.

Teorema 1.3.2 *Seja* $J = \operatorname{diag}(1, -1)$, $C = \operatorname{diag}(c_1, c_2) \in M_2$ *e* $A = [a_{ij}] \in M_2$, *com* $\sigma(A) = \{\alpha_1, \alpha_2\}$, $\alpha = \frac{1}{2}\operatorname{Tr}(A)\operatorname{Tr}(C)$. *Sejam*

$$\begin{aligned} M &= (c_1 - c_2)^2\left(|\alpha_1|^2 + |\alpha_2|^2 - \operatorname{Tr}\big(A^{[*]}A\big)\right), \\ N &= (c_1 - c_2)^2\left(\operatorname{Tr}\big(A^{[*]}A\big) - 2\operatorname{Re}(\bar{\alpha}_1\alpha_2)\right). \end{aligned}$$

Consideram-se ainda os pontos-σ $z_{\text{id}} = c_1\alpha_1 + c_2\alpha_2$ *e* $z_{(12)} = c_1\alpha_2 + c_2\alpha_1$.

a) *Se* $M > 0$ *e* $N > 0$, *então* $W_C^J(A)$ *é limitado por um dos ramos da hipérbole de focos* z_{id} *e* $z_{(12)}$, *cujos eixos transverso e não-transverso têm comprimento* $\sqrt{N}$ *e* $\sqrt{M}$, *respectivamente.*

b) *Se* $M > 0$ *e* $N = 0$, *então* $W_C^J(A)$ *é:*

- **i.** *a recta* l, *se* $|a_{12}| = |a_{21}|$;
- **ii.** *um semi-plano aberto definido pela recta* l, *se* $|a_{12}| \neq |a_{21}|$;

em que l *é a recta que passa por* α *e é perpendicular ao segmento que une* z_{id} *a* $z_{(12)}$.

c) *Se* $M > 0$ *e* $N < 0$, *então* $W_C^J(A)$ *é todo o plano complexo.*

d) *Se* $M = 0$ *e* $N > 0$, *então* $W_C^J(A)$ *é uma semi-recta fechada, sobre a recta definida por* z_{id} *e* $z_{(12)}$, *com um dos pontos-*σ *por extremo.*

e) *Se* $M = N = 0$, *então* $W_C^J(A)$ *é:*

i. *o conjunto singular* $\{\alpha\}$*, se* $a_{11} = a_{22}$ *ou* $c_1 = c_2$*;*

ii. *uma semi-recta aberta, sobre a recta definida por* $c_1 a_{11} + c_2 a_{22}$ *e* $c_2 a_{11} + c_1 a_{22}$*, de extremo* α*, se* $a_{11} \neq a_{22}$ *e* $c_1 \neq c_2$*.*

Demonstração: Como $C = c_2\, I_2 + (c_1 - c_2)\, E_{11}$ e $W^J_{E_{11}}(A) = W^+_J(A)$, pois $j_1 = 1$, pelas propriedades $\mathrm{W}^{\mathrm{H}}_{\mathrm{C}}2$ e $\mathrm{W}^{\mathrm{H}}_{\mathrm{C}}4$, tem-se

$$W^J_C(A) \;=\; c_2 \mathrm{Tr}(A) + (c_1 - c_2) W^+_J(A). \tag{1.53}$$

Se C é uma matriz escalar, é óbvio que $W^J_C(A) = \{\alpha\}$. Se C é não-escalar, o resultado é consequência de (1.53) e do Teorema 1.2.7. ■

Corolário 1.3.1 *Se* $J = \mathrm{diag}\,(1, -1)$*,* $C = \mathrm{diag}\,(c_1, c_2) \in M_2(\mathbb{R})$ *e* $A = [a_{ij}] \in M_2$ *é uma matriz Hermítica-J de valores próprios* α_1 *e* α_2*, tal que* $z_{\mathrm{id}} = c_1\alpha_1 + c_2\alpha_2$*, então* $W^J_C(A)$ *é:*

i) $\mathbb{R}$*, se* $\alpha_1,\, \alpha_2 \in \mathbb{C} \setminus \mathbb{R}$*;*

ii) $[z_{\mathrm{id}}, +\infty)$*, se* $(\alpha_1 > \alpha_2$ *e* $a_{11} > a_{22}$ *e* $c_1 > c_2)$ *ou* $(\alpha_1 < \alpha_2$ *e* $a_{11} < a_{22}$ *e* $c_1 < c_2)$*;*

iii) $(-\infty, z_{\mathrm{id}}]$*, se* $(\alpha_1 < \alpha_2$ *e* $a_{11} < a_{22}$ *e* $c_1 > c_2)$ *ou* $(\alpha_1 > \alpha_2$ *e* $a_{11} > a_{22}$ *e* $c_1 < c_2)$*;*

iv) $\{z_{\mathrm{id}}\}$*, se* $(\alpha_1 = \alpha_2$ *e* $a_{11} = a_{22})$ *ou* $c_1 = c_2$*, isto é,* $A = \alpha_1 I_2$ *ou* $C = c_1 I_2$*;*

v) $(z_{\mathrm{id}}, +\infty)$*, se* $\alpha_1 = \alpha_2$ *e* $(c_1 - c_2)(a_{11} - a_{22}) > 0$*;*

vi) $(-\infty, z_{\mathrm{id}})$*, se* $\alpha_1 = \alpha_2$ *e* $(c_1 - c_2)(a_{11} - a_{22}) < 0$*;*

com $\alpha_1, \alpha_2 \in \mathbb{R}$ *nos casos* ii) - vi).

Demonstração: Atendendo a (1.53), o resultado segue facilmente do Corolário 1.2.4 que apresenta a descrição minuciosa de $W^+_J(A)$, quando A é Hermítica-J. ■

A análise atenta e criteriosa do caso 2×2 permite derivar outros resultados interessantes para o contradomínio numérico, para além da convexidade, como a caracterização das matrizes cujo contradomínio numérico é um conjunto singular ou o subconjunto de uma recta ou um polígono (e seu interior), e a caracterização de pontos da fronteira não diferenciáveis. Dirigimos a atenção destes problemas para o contradomínio tracial-C, J.

Condições para $W^J_C(A)$ ser um Conjunto Singular

Dada $A \in M_n$ e $k, l \in \{1, \ldots, n\}$, denota-se por $A[kl]$ a submatriz principal de A determinada pelas linhas e colunas k e l.

Lema 1.3.1 *Sejam* $C = \operatorname{diag}(c_1, \ldots, c_n)$, $A = [a_{ij}] \in M_n$. *Então* $W_C^J(A)$ *contém o conjunto*

$$W_{C_{kl}}^{J_{kl}}(A_{kl}) + \sum_{i \neq k,l} c_i a_{ii}, \tag{1.54}$$

onde $A_{kl} = A[kl]$, $C_{kl} = C[kl]$ *e* $J_{kl} = J[kl]$, $1 \leq k < l \leq n$.

Demonstração: Seja z um elemento arbitrário do conjunto (1.54), isto é,

$$z = \operatorname{Tr}(C_{kl} M^{-1} A_{kl} M) + \sum_{i \neq k,l} c_i a_{ii},$$

em que $M \in M_2$ é uma matriz satisfazendo $M^* J_{kl} M = J_{kl}$. Sejam

$$A' = \begin{bmatrix} A_{kl} & A_{12} \\ A_{21} & \check{A}_{kl} \end{bmatrix}, \qquad C' = \begin{bmatrix} C_{kl} & 0 \\ 0 & \check{C}_{kl} \end{bmatrix},$$

e U_M as matrizes que se obtêm de A, C e $M \oplus I_{n-2}$, respectivamente, permutando as linhas e colunas 1 e 2 pelas linhas e colunas k e l. Observa-se que

$$(M \oplus I_{n-2})^{-1} A' (M \oplus I_{n-2}) = \begin{bmatrix} M^{-1} A_{kl} M & M^{-1} A_{12} \\ A_{21} M & \check{A}_{kl} \end{bmatrix}.$$

Mediante cálculos simples, verifica-se que

$$z = \operatorname{Tr}(C_{kl} M^{-1} A_{kl} M) + \operatorname{Tr}(\check{C}_{kl} \check{A}_{kl}) = \operatorname{Tr}(C'(M \oplus I_{n-2})^{-1} A'(M \oplus I_{n-2})) = \operatorname{Tr}(C U_M^{[*]} A U_M),$$

onde $U_M \in M_n$ satisfaz $U_M^* J U_M = J$. Portanto, $z \in W_C^J(A)$. ■

Se $A \in M_n$ ou $C \in M_n$ é uma matriz escalar, é claro que $W_C^H(A) = \left\{ \frac{1}{n} \operatorname{Tr}(A) \operatorname{Tr}(C) \right\}$.

Teorema 1.3.3 *Seja* $C \in M_n$ *uma matriz diagonal não-escalar. Dada* $A \in M_n$, $W_C^J(A)$ *é um conjunto singular se e só se* A *é uma matriz escalar.*

Demonstração: A implicação ($\Leftarrow$) é óbvia.

($\Rightarrow$) Seja $C_\sigma = P_\sigma C P_\sigma^T$, em que P_σ é a matriz de permutação associada a $\sigma \in S_n$ e $C = \operatorname{diag}(c_1, \ldots, c_n)$. Suponhamos que $A = [a_{ij}]$ é uma matriz não-escalar. É possível encontrar k, l, tais que $1 \leq k < l \leq n$ (com $k \leq r < l$, se $0 < r < n$), e $\sigma = \sigma_1 \sigma_2$, com σ_1 e σ_2 permutações de $1, \ldots, r$ e $r+1, \ldots, n$, respectivamente, de tal modo que $A_{kl} = A[kl]$ e $C'_{kl} = C_\sigma[kl]$ são submatrizes principais não-escalares de A e C_σ, respectivamente. Nestas condições, a adjunta-J de P_σ é P_σ^T e, pela propriedade $\mathrm{W}_\mathrm{C}^\mathrm{H}1$, tem-se $W_C^J(A) = W_{C_\sigma}^J(A)$. Pelo Lema 1.3.1, $W_{C_\sigma}^J(A)$ contém o conjunto

$$\Gamma_{kl} = W_{C'_{kl}}^{J_{kl}}(A_{kl}) + \sum_{i \neq k,l} c_{\sigma(i)} a_{ii}.$$

Se $r = 0$ ou $r = n$, então $W^{J_{kl}}_{C'_{kl}}(A_{kl}) = W_{C'_{kl}}(A_{kl})$ e, pelo Teorema 1.3.1, o subconjunto Γ_{kl} de $W^J_C(A)$ é um disco elíptico, eventualmente degenerado, mas nunca um ponto. Se $0 < r < n$, então $J_{kl} = \mathrm{diag}\,(1, -1)$ e, pelo Teorema 1.3.2, o subconjunto Γ_{kl} de $W^J_C(A)$ é um ramo de uma hipérbole com interior, possivelmente degenerada, mas nunca um ponto. Em qualquer caso, $W^J_C(A)$ não é um conjunto singular. ■

Condições para $W^J_C(A)$ ser um Subconjunto de uma Recta

Agora, generalizamos o resultado de $W_C(A)$ ser um subconjunto do eixo real se e só se A é uma matriz Hermítica, quando $C \in M_n(\mathbb{R})$ é diagonal, ao contradomínio tracial-C, J.

Teorema 1.3.4 *Seja $C \in M_n(\mathbb{R})$ uma matriz diagonal não-escalar. Dada $A \in M_n$, $W^J_C(A)$ é um subconjunto do eixo real se e só se A é uma matriz Hermítica-J.*

Demonstração: ($\Leftarrow$) Qualquer $z \in W^J_C(A)$ é da forma $z = \mathrm{Tr}\left(CU^{[*]}AU\right)$, $U \in U_{r,n-r}$, e satisfaz

$$\bar{z} \;=\; \mathrm{Tr}\left(CU^{[*]}AU\right)^* \;=\; \mathrm{Tr}\left(JU^{[*]}A^{[*]}UJC\right).$$

Como C é diagonal, é claro que $JCJ = C$. Se $A^{[*]} = A$, então $\bar{z} = z$, logo $W^J_C(A) \subseteq \mathbb{R}$.

($\Rightarrow$) Supondo que $A = [a_{ij}]$ não é Hermítica-J, então A contém uma submatriz principal $A_{kl} = A[kl]$, $1 \le k < l \le n$ (com $k \le r < l$, se $0 < r < n$), que não é Hermítica-J_{kl} onde $J_{kl} = J[kl]$. Como na demonstração do Teorema 1.3.3, atendendo de novo ao Lema 1.3.1, $W^J_C(A)$ contém o conjunto

$$\Gamma_{kl} \;=\; W^{J_{kl}}_{C'_{kl}}(A_{kl}) \;+\; \sum_{i \ne k,l} c_{\sigma(i)} a_{ii},$$

onde $\sigma = \sigma_1\sigma_2$, sendo σ_1 e σ_2 permutações de $1, \ldots, r$ e $r+1, \ldots, n$, respectivamente, tal que $C'_{kl} = C_\sigma[kl]$ é uma submatriz principal não escalar de $C_\sigma = P_\sigma C P_\sigma^T$. Repetindo os passos da demonstração do Teorema 1.3.3, o subconjunto Γ_{kl} de $W^J_C(A)$ é um disco elíptico ou um ramo de uma hipérbole com interior, que nunca degeneram num subconjunto do eixo real. ■

O Corolário que se segue é uma consequência imediata do Teorema 1.3.4.

Corolário 1.3.2 *Seja $C \in M_n(\mathbb{R})$ uma matriz diagonal não-escalar. Dada $A \in M_n$, $W^J_C(A)$ é um subconjunto de uma linha recta se e só se A é uma matriz essencialmente Hermítica-J.*

Demonstração: ($\Leftarrow$) Da definição de matriz essencialmente Hermítica-J, pela propriedade $\mathrm{W}_{\mathrm{C}}^{\mathrm{H}}2$ e pelo Teorema 1.3.4, tem-se o pretendido.
($\Rightarrow$) Se $W_C^J(A)$ é um subconjunto de uma linha recta, pela propriedade $\mathrm{W}_{\mathrm{C}}^{\mathrm{H}}2$, pode aplicar-se uma rotação e uma translação a $W_C^J(A)$, de modo a obter-se um subconjunto do eixo real, isto é, existem escalares $\alpha, \beta \in \mathbb{C}$, $\alpha \neq 0$, tais que $W_C^J(\alpha I_n + \beta A) \subseteq \mathbb{R}$. Pelo Teorema 1.3.4, a matriz $\alpha I_n + \beta A$ é Hermítica-J, concluindo-se que A é essencialmente Hermítica-J. ■

1.3.2 Uma Consequência do Teorema de Tarski

Recorrendo a um teorema devido a Tarski, Leal Duarte [48] mostrou que a fronteira de $W_C(A)$, $A, C \in M_n$, é uma união finita de arcos algébricos. Nakazato, Mishikawa e Takaguchi [108] seguiram outra abordagem e obtiveram o mesmo resultado. Provaremos que um resultado análogo vale para o contradomínio tracial-C, H.

Após introduzir alguma notação útil, enunciamos o Teorema de Tarski, remetendo a sua demonstração para [83]. Denota-se por $\mathbb{Z}[t_1, \ldots, t_r; x_1, \ldots, x_s]$ o anel de polinómios sobre o conjunto dos números inteiros, constituído pelos polinómios $f(t_1, \ldots, t_r; x_1, \ldots, x_s)$ nas variáveis $x_1, \ldots, x_s$ e de coeficientes em $\mathbb{Z}[t_1, \ldots, t_r]$.

Teorema 1.3.5 (Teorema de Tarski) *Sejam $f_1, \ldots, f_m$ e $g_1, \ldots, g_n$ polinómios no anel $\mathbb{Z}[t_1, \ldots, t_r; x_1, \ldots, x_s]$. É possível encontrar, num número finito de passos, uma colecção finita de conjuntos $\psi_1, \ldots, \psi_p$, onde*

$$\psi_l = \{F_{l_1}, \ldots, F_{l_{p_l}}; G_{l_1}, \ldots, G_{l_{q_l}}\}, \quad F_{l_i}, G_{l_j} \in \mathbb{Z}[t_1, \ldots, t_r], \quad 1 \leq i \leq p_l,\ 1 \leq j \leq q_l,$$

tal que, qualquer que seja $w = (w_1, \ldots, w_r) \in R$, em que R é um corpo real fechado arbitrário, as equações polinomiais

$$\begin{array}{lll} f_1(w_1, \ldots, w_r; x_1, \ldots, x_s) = 0, & \ldots, & f_m(w_1, \ldots, w_r; x_1, \ldots, x_s) = 0, \\ g_1(w_1, \ldots, w_r; x_1, \ldots, x_s) > 0, & \ldots, & g_n(w_1, \ldots, w_r; x_1, \ldots, x_s) > 0 \end{array}$$

têm uma solução se e só se existe pelo menos um l, $1 \leq l \leq p$, tal que

$$F_{l_i}(w) = 0, \quad 1 \leq i \leq p_l, \qquad G_{l_j}(w) > 0, \quad 1 \leq j \leq q_l,$$

também tem uma solução.

Seguindo de perto a abordagem adoptada por Leal Duarte [48], aplica-se o Teorema de Tarski para provar que a fronteira de $W_C^H(A)$ é constituída por uma união finita de arcos algébricos.

Teorema 1.3.6 *Sejam $A, C \in M_n$ matrizes diagonais. A fronteira de $W_C^H(A)$ é uma união finita de arcos algébricos e, portanto, uma curva de classe $\mathbb{C}^\infty$, excepto possivelmente num número finito de pontos.*

Demonstração: Primeiro, mostra-se que estamos nas condições de aplicar o Teorema de Tarski. Sejam $A = \operatorname{diag}(\alpha_1, \ldots, \alpha_n)$ e $C = \operatorname{diag}(c_1, \ldots, c_n)$. Ora, $z \in W_C^H(A)$ se e só se existe uma matriz U unitária-H tal que $z = \operatorname{Tr}(CU^{[*]}AU)$. Sejam $H = [a_{kj} + ib_{kj}]$ e $U = [x_{kj} + iy_{kj}]$, onde $a_{kj}, b_{kj}, x_{kj}, y_{kj} \in \mathbb{R}$, $k, j = 1, \ldots, n$. Os pontos de $W_C^H(A)$ satisfazem as seguintes condições nas variáveis x_{kj} e y_{kj}:

$$\operatorname{Re} z + i \operatorname{Im} z - \operatorname{Tr}(C[a_{kj} + ib_{kj}]^{-1}[x_{kj} + iy_{kj}]^*[a_{kj} + ib_{kj}]A[x_{kj} + iy_{kj}]) \;=\; 0, \quad (1.55)$$

$$\sum_{k,l=1}^{n} (x_{km} - i\, y_{km})(a_{kl} + i\, b_{kl})(x_{lr} + i\, y_{lr}) - (a_{mr} + i\, b_{mr}) \;=\; 0, \quad 1 \leq m, r \leq n. \quad (1.56)$$

Os primeiros membros das equações (1.55) e (1.56) são polinómios nos parâmetros $\operatorname{Re} z$, $\operatorname{Im} z$, $\operatorname{Re} \alpha_h$, $\operatorname{Im} \alpha_h$, $\operatorname{Re} c_h$, $\operatorname{Im} c_h$, $h = 1, \ldots, n$, a_{kj}, b_{kj}, $k, j = 1, \ldots, n$, e nas variáveis x_{kj}, y_{kj}, $k, j = 1, \ldots, n$. Considerando as partes reais e imaginárias destes polinónios e igualando-as a zero, obtém-se um conjunto de equações polinomiais de coeficientes inteiros nos mesmos parâmetros e variáveis. Pelo Teorema de Tarski, pode concluir-se que existe uma colecção finita de conjuntos $\psi_1, \ldots, \psi_p$, onde

$$\psi_l = \{F_{l_1}, \ldots, F_{l_{n_l}}; G_{l_1}, \ldots, G_{l_{m_l}}\}, \;\; F_{l_s}, G_{l_q} \in \mathbb{Z}[t_1, \ldots, t_{n^2+4n+2}], \;\; 1 \leq s \leq n_l, \;\; 1 \leq q \leq m_l,$$

tais que, para qualquer

$$w \;=\; ((\operatorname{Re} z, \operatorname{Im} z), (\operatorname{Re} \alpha_h, \operatorname{Im} \alpha_h, \operatorname{Re} c_h, \operatorname{Im} c_h,\; h = 1, \ldots, n), (a_{kj}, b_{kj},\; k, j = 1, \ldots, n)),$$

em $\mathbb{C}^{n^2+4n+2}$, as equações iniciais (1.55) e (1.56) têm uma solução se e só se existe pelo menos um l, $1 \leq l \leq p$, tal que

$$F_{l_s}(w) \;=\; 0, \quad 1 \leq s \leq n_l, \qquad G_{l_q}(w) \;>\; 0, \quad 1 \leq q \leq m_l, \quad (1.57)$$

tem uma solução. Um ponto que apenas satisfaça as desigualdade em (1.57) (e nenhuma igualdade) não pertence à fronteira de $W_C^H(A)$. A fronteira de $W_C^H(A)$ é constituída por um número finito de arcos algébricos. Os pontos fronteiros são de classe $\mathbb{C}^\infty$, a menos que satisfaçam as equações (1.57) para, no mínimo, dois valores de l, o que apenas se verifica para um número finito de pontos [83, p.325]. ■

O resultado anterior pode ser generalizado a matrizes arbitrárias $A, C \in M_n$.

Aplicando este resultado a $C = E_{11}$ (E_{nn}), com $0 < r < n$, conclui-se que a fronteira de $W_C^J(A) = W_J^+(A)$ $(-W_J^-(A))$ é uma união finita de arcos algébricos. Em particular, para $A_R = R^{-1}AR$, em que R é uma matriz não-singular tal que $R^*HR = J$ é a matriz de inércia de H, conclui-se que as curvas geradoras de fronteira de $W_H(A) = W_J(A_R)$ são curvas algébricas, respondendo afirmativamente a uma questão colocada por C.-K. Li, N. K. Tsing e F. Uhlig [95, p.16].

1.3.3 Pontos Angulosos do Contradomínio Tracial-C, J

Consideremos $J = I_r \oplus -I_{n-r}$, $0 \leq r \leq n$, e $C \in M_n$ uma matriz diagonal. O teorema seguinte, sobre os pontos angulosos de $W_C^J(A)$, generaliza um importante resultado de Natália Bebiano [11], relativo ao contradomínio numérico-C.

Teorema 1.3.7 *Seja $C = c_1 I_{n_1} \oplus \cdots \oplus c_p I_{n_p}$, $n_1, \ldots, n_p \in \mathbb{N}$, $n_1 + \cdots + n_p = n$, com $c_1, \ldots, c_p$ números complexos distintos dois a dois. Se $A \in M_n$ e $z = \operatorname{Tr}\left(CU^{[*]}AU\right)$, $U \in U_{r,n-r}$, é um ponto anguloso de $W_C^J(A)$, então*

$$U^{[*]}AU \;=\; A_1 \oplus \cdots \oplus A_p \qquad e \qquad z \;=\; \sum_{i=1}^{p} c_i \operatorname{Tr}(A_i), \qquad A_i \in M_{n_i}, \quad i = 1, \ldots, p.$$

Demonstração: Para simplificar, escreva-se $A_U = U^{[*]}AU$. Como $J = I_r \oplus -I_{n-r}$, tem-se $J^{1/2} = I_r \oplus i\, I_{n-r}$. Dada $S \in M_n$ uma matriz Hermítica arbitrária e $t \in \mathbb{R}$ numa vizinhança de zero,

$$\mathrm{e}^{itJ^{1/2}SJ^{1/2}} \;=\; I + itJ^{1/2}SJ^{1/2} + O(t^2)$$

é uma matriz pseudo-unitária em $U_{r,n-r}$. Considera-se a função diferenciável

$$f(t) \;=\; \operatorname{Tr}\left(C\,\mathrm{e}^{-itJ^{1/2}SJ^{1/2}} A_U\, \mathrm{e}^{itJ^{1/2}SJ^{1/2}}\right).$$

Como $z = \operatorname{Tr}(CA_U)$ é um ponto anguloso de $W_C^J(A)$, a derivada de f em relação a t no ponto $t = 0$ é nula. Ora

$$f'(0) \;=\; i \operatorname{Tr}\left(CA_U J^{1/2}SJ^{1/2} - CJ^{1/2}SJ^{1/2}A_U\right) \;=\; i \operatorname{Tr}\left(SJ^{1/2}[C, A_U]J^{1/2}\right),$$

dada a propriedade cíclica do traço e usando a notação usual de comutadores de matrizes. Sendo S arbitrária, de $f'(0) = 0$, conclui-se que $J^{1/2}[C, A_U]J^{1/2} = 0$ e, assim,

$$[C, A_U] \;=\; 0. \tag{1.58}$$

Por hipótese, C é uma soma directa de p matrizes escalares $c_i I_{n_i}$, com c_i, $i = 1, \ldots, p$, todos distintos. Portanto, segue de (1.58), ou seja, do facto de A_U e C comutarem entre si, que A_U é uma soma directa do tipo $A_1 \oplus \cdots \oplus A_p$, em que $A_i \in M_{n_i}$, $i = 1, \ldots, p$. Verifica-se, com facilidade, que

$$z \;=\; \operatorname{Tr}(CA_U) \;=\; \sum_{i=1}^{p} c_i \operatorname{Tr}(A_i). \qquad \blacksquare$$

Apresentam-se algumas consequências imediatas do Teorema 1.3.7.

Corolário 1.3.3 *Seja $C = I_k \oplus 0_{n-k}$, $1 \leq k \leq n$. Se $A \in M_n$ e $z \in W_C^J(A)$ é um ponto anguloso de $W_C^J(A)$, então existe $U \in U_{r,n-r}$ tal que*

$$U^{[*]}AU \;=\; A_1 \oplus A_2 \qquad e \qquad z \;=\; \mathrm{Tr}(A_1), \qquad A_1 \in M_k.$$

Se $J = I_n$, então o Teorema 1.3.7 e o Corolário 1.3.3 caracterizam os pontos angulosos do contradomínio numérico-C e do contradomínio numérico-k de A, respectivamente.

Corolário 1.3.4 *Seja $A \in M_n$. Se z é um ponto anguloso do conjunto $\pm W_J^{\pm}(A)$, então $z \in \sigma_J^{\pm}(A)$ e existe $x \in \mathbb{C}^n$, tal que*

$$Ax \;=\; zx, \qquad A^{[*]}x \;=\; \bar{z}x \qquad e \qquad \langle x, x\rangle_J = \pm 1.$$

Demonstração: Seja $z \in W_J^+(A)$ um ponto anguloso de $W_J^+(A)$. Então $J \neq -I_n$ e, pelo Corolário 1.3.3, com $k = 1$, existe $U \in U_{r,n-r}$, $r \neq 0$, tal que $U^{[*]}AU = [z] \oplus A_2$, onde $A_2 \in M_{n-1}$. Então $U^{[*]}A^{[*]}U = [\bar{z}] \oplus B_2$, em que B_2 é a adjunta-J[11] de A_2. Obtém-se, assim, $U^{[*]}AUe_1 = ze_1$ e $U^{[*]}A^{[*]}Ue_1 = \bar{z}e_1$, onde e_1 é o primeiro vector da base canónica de $\mathbb{C}^n$. Como $UU^{[*]} = I_n$, tem-se $AUe_1 = zUe_1$ e $A^{[*]}Ue_1 = \bar{z}Ue_1$, ou seja, $x = Ue_1$ é simultaneamente um vector próprio de A associado ao valor próprio z e um vector próprio de $A^{[*]}$ associado ao valor próprio $\bar{z}$, satisfazendo $\langle x, x\rangle_J = \langle e_1, e_1\rangle_J = 1$. De modo análogo, prova-se o resultado quando z é um ponto anguloso de $-W_J^-(A)$. ■

Nota 1.3.1 O Lema 1.2.1, embora mais geral, é consequência imediata do Corolário 1.3.4. De facto, seja $z \in \pm W_H^{\pm}(A)$ um ponto anguloso de $\pm W_H^{\pm}(A)$. Ora, dada $A_R = R^{-1}AR$, onde R é uma matriz não-singular tal que $R^*HR = J$ é a matriz de inércia de H, tem-se $W_H^{\pm}(A) = W_J^{\pm}(A_R)$. Pelo Corolário 1.3.4 aplicado à matriz A_R, existe $x \in \mathbb{C}^n$, tal que $A_Rx = zx$, $JA_R^*Jx = \bar{z}x$ e $\langle x, x\rangle_J = \pm 1$. Donde resulta, para $y = Rx$, que $Ay = zy$, $H^{-1}A^*Hy = zy$ e $\langle y, y\rangle_H = \langle x, x\rangle_J = \pm 1$, portanto, $z \in \sigma_H^{\pm}(A)$ e $\bar{z} \in \sigma_H^{\pm}(A^{[*]})$.

Se A e C são matrizes nas hipóteses do Teorema 1.3.7, é óbvio que $\det(A - \lambda I_n) = 0$ se e só se $\det(A_i - \lambda I_{n_i}) = 0$, $i = 1, \ldots, p$. Portanto, os pontos angulosos de $W_C^J(A)$ são do tipo

$$z \;=\; \sum_{i=1}^{p} c_i \sum_{k=1}^{n_i} \alpha_k^i,$$

em que α_k^i, $k = 1, \ldots, n_i$, são os valores próprios da matriz A associados ao bloco A_i, $i = 1, \ldots, p$. O resultado seguinte é, assim, uma consequência trivial do Teorema 1.3.7, quando $p = n$, e relaciona os pontos angulosos e os pontos-σ do contradomínio tracial-C, J.

Corolário 1.3.5 *Seja $C = \operatorname{diag}(c_1, \ldots, c_n) \in M_n$, com $c_1, \ldots, c_n$ todos distintos. Se $A \in M_n$ tem valores próprios $\alpha_1, \ldots, \alpha_n$ e $z \in W_C^J(A)$ é um ponto anguloso de $W_C^J(A)$, então z é um ponto-σ do tipo*

$$z_\sigma = \sum_{i=1}^{r} c_i \alpha_{\sigma_1(i)} + \sum_{i=r+1}^{n} c_i \alpha_{\sigma_2(i)}, \tag{1.59}$$

com $\sigma = \sigma_1\sigma_2 \in S_n$, para σ_1 e σ_2 permutações de $1, \ldots, r$ e $r+1, \ldots, n$, respectivamente.

Demonstração: Pelo Teorema 1.3.7, com $p = n$, existe uma matriz pseudo-unitária $U \in U_{r,n-r}$, tal que $U^{[*]}AU$ é um matriz diagonal, cujas entradas da diagonal principal são os valores próprios de A. Assim, $U^{[*]}AU = P_\sigma^T \operatorname{diag}(\alpha_1, \ldots, \alpha_n) P_\sigma$, onde P_σ é a matriz de permutação associada à permutação σ, e resulta da observação que precedeu este corolário que z é um ponto-σ do tipo (1.59). ∎

A análise do caso 2×2 é ainda útil no estudo de determinados pontos especiais da fronteira do contradomínio tracial-C, J.

Teorema 1.3.8 *Sejam $C = \operatorname{diag}(c_1, \ldots, c_n) \in M_n$ e $A = [a_{ij}] \in M_n$ uma matriz triangular superior (inferior). Se $\operatorname{Tr}(CA) \in \partial W_C^J(A)$ e $c_k \neq c_l$, então $a_{kl} = 0$, para $1 \leq k < l \leq n$.*

Demonstração: Suponhamos que A é triangular superior. Sejam $1 \leq k < l \leq n$, tais que $c_k \neq c_l$, e considere-se a submatriz principal de A, determinada pelas linhas e colunas k e l,

$$A_{kl} = \begin{bmatrix} a_{kk} & a_{kl} \\ 0 & a_{ll} \end{bmatrix},$$

com $a_{kl} \neq 0$. Pelo Lema 1.3.1, $W_C^J(A)$ contém o subconjunto

$$\Gamma_{kl} = W_{C_{kl}}^{J_{kl}}(A_{kl}) - c_k a_{kk} - c_l a_{ll} + \operatorname{Tr}(CA)$$

onde $C_{kl} = C[kl]$ e $J_{kl} = J[kl]$. O conjunto $W_{C_{kl}}^{J_{kl}}(A_{kl})$ é um disco elíptico (se $j_k = j_l$) ou um ramo de uma hipérbole com interior (se $j_k = -j_l$), possivelmente degenerados. Em ambos os casos, $c_k a_{kk} + c_l a_{ll}$ é um dos focos desta cónica. Como A_{kl} não é essencialmente Hermítica-J_{kl} (caso contrário $a_{kl} = 0$) e C_{kl} não é escalar, $W_{C_{kl}}^{J_{kl}}(A_{kl})$ pode, quando muito, reduzir-se a um disco circular, a um semi-plano ou a todo o plano, mas não a um subconjunto de uma linha recta. Tal significa que $c_k a_{kk} + c_l a_{ll}$ é um ponto interior a $W_{C_{kl}}^{J_{kl}}(A_{kl})$. Daqui se conclui que $\operatorname{Tr}(CA)$ é um ponto interior ao subconjunto Γ_{kl} de $W_C^J(A)$, o que contradiz a hipótese de pertencer a $\partial W_C^J(A)$, a menos que a_{kl} se anule. Análogo racíocinio vale para A triangular inferior. ∎

Os dois últimos corolários desta secção são consequência óbvia do Teorema 1.3.8.

Corolário 1.3.6 *Seja $C = c_1 I_{n_1} \oplus \cdots \oplus c_p I_{n_p}$, $n_1, \ldots, n_p \in \mathbb{N}$, $n_1 + \cdots + n_p = n$, com $c_1, \ldots, c_p$ números complexos todos distintos. Se $A \in M_n$ é uma matriz triangular superior (inferior) e $\mathrm{Tr}(CA) \in \partial W_C^J(A)$, então $A = A_1 \oplus \cdots \oplus A_p$, onde cada $A_i \in M_{n_i}$ é uma matriz triangular superior (inferior), $i = 1, \ldots, p$.*

Corolário 1.3.7 *Seja $C = \mathrm{diag}\,(c_1, \ldots, c_n) \in M_n$, em que $c_1, \ldots, c_p$ são dois a dois distintos. Se $A \in M_n$ é uma matriz triangular superior (inferior) e $\mathrm{Tr}(CA) \in \partial W_C^J(A)$, então A é uma matriz diagonal.*

1.4 Contradomínio Determinantal-C, H

Dada $C \in M_n$, o *contradomínio determinantal-C* de $A \in M_n$ define-se por

$$\triangle_C(A) \;=\; \{\det(C + U^*AU) :\; U \in M_n,\; U^*U = I_n\}\,.$$

Sendo a imagem da órbita unitária $\mathcal{O}(A)$ pela aplicação contínua $X \mapsto \det(C + X)$, pode entender-se $\triangle_C(A)$ como uma variação do contradomínio numérico-C. Aliás, um interessante paralelismo existe entre estes dois conjuntos, que partilham propriedades similares [10, 12].

A primeira propriedade em comum é obviamente o facto de $W_C(A)$ e $\triangle_C(A)$ serem conjuntos conexos e compactos, para quaisquer $A, C \in M_n$. Ambos os conjuntos são invariantes por transformações de semelhança unitária de A ou de C; e reduzem-se a um ponto, quando uma das matrizes A ou C é escalar.

Definem-se pontos-σ de $\triangle_C(A)$ por

$$z_\sigma \;=\; \prod_{i=1}^{n} \left(c_i + \alpha_{\sigma(i)}\right), \qquad \sigma \in S_n, \tag{1.60}$$

onde α_i e c_i, $i = 1, \ldots, n$, são os valores próprios de A e C, respectivamente. Os pontos-σ (1.60) pertencem a $\triangle_C(A)$ e, tal como os pontos-σ de $W_C(A)$, o seu papel é de destaque. Por exemplo, se $A, C \in M_n$ são ambas Hermíticas, então os extremos dos conjuntos compactos reais $W_C(A)$ e $\triangle_C(A)$ são pontos-σ [104, 52].

Apesar das propriedades semelhantes encontradas, $W_C(A)$ e $\triangle_C(A)$ exibem diferenças cruciais. O conjunto $\triangle_C(A)$ é estrelado para $A, C \in M_3$ normais [14], contudo $\triangle_C(A)$ pode nem sequer ser simplesmente conexo [10], ainda que A, C sejam Hermíticas. De modo geral, o grau de dificuldade dos problemas envolvendo $\triangle_C(A)$ supera largamente os correspondentes problemas envolvendo o conjunto $W_C(A)$.

Com vista a documentar a afirmação anterior, reportamo-nos a uma famosa conjectura, relativa ao determinante da soma de duas matrizes normais com valores próprios prescritos. Devida a Marvin Marcus [100] e Graciano N. de Oliveira [113], pode ser reformulada, em termos do contradomínio determinantal-C e seus pontos-σ, pela seguinte inclusão:

$$\triangle_C(A) \;\subseteq\; \mathrm{conv}\,\{z_\sigma : \sigma \in S_n\}\,, \tag{1.61}$$

para $A, C \in M_n$ matrizes normais. Este problema continua em aberto para $n \geq 4$, com excepção de alguns casos particulares. A conjectura encontra-se resolvida, por exemplo, quando os pontos-σ de $\triangle_C(A)$ são colineares; se $\det(A + C) = 0$; quando A é definida positiva e $\sigma(C) \subseteq i\mathbb{R}$; se $\sigma(A) \cap \sigma(C)$ contém um valor próprio de multiplicidade não inferior a $n - 2$; se A ou C é essencialmente Hermítica. Os principais avanços obtidos na área devem-se a N. Bebiano, S. W. Drury, A. Kovačec, J. K. Merikoski, J. da Providência e A. Virtanen (alguns resultados podem ser revistos em [16]).

A solução do problema "dual", relativo à inclusão de $W_C(A)$ no invólucro convexo dos seus pontos-σ é conhecida [113]. Se os valores próprios da matriz normal $C \in M_n$ são colineares, da convexidade de $W_C(A)$, tem-se a igualdade

$$W_C(A) \;=\; \mathrm{conv}\,\{z_\sigma : \sigma \in S_n\}\,,$$

para qualquer $A \in M_n$ normal. Em virtude de $\triangle_C(A)$ não ser, em geral, convexo, mesmo que a inclusão (1.61) seja válida, esta não se converte em igualdade.

Perspectivando-se a conjectura de Marcus-Oliveira de elevada complexidade, fica a certeza das potencialidades decorrentes de explorar as propriedades de $\triangle_C(A)$.

Como variação do contradomínio tracial-C, H apresenta-se o *contradomínio determinantal-C, H* de $A \in M_n$ que se define por

$$\triangle_C^H(A) \;=\; \left\{\det\left(C + U^{[*]}AU\right) :\; U \in M_n,\; U^*HU = H\right\}.$$

Se $H = \pm I_n$, este conceito reduz-se a $\triangle_C(A)$.

Listam-se algumas propriedades básicas do contradomínio determinantal-C, H que podem ser entendidas como duais das propriedades $\mathrm{W}_\mathrm{C}^\mathrm{H}1$, $\mathrm{W}_\mathrm{C}^\mathrm{H}3$ e $\mathrm{W}_\mathrm{C}^\mathrm{H}4$ do contradomínio tracial-C, H :

$\mathbf{D}_\mathbf{C}^\mathbf{H}\mathbf{1}$. $\triangle^H_{V^{[*]}CV}(U^{[*]}AU) = \triangle_C^H(A)$, para quaisquer matrizes U e V unitárias-H;

$\mathbf{D}_\mathbf{C}^\mathbf{H}\mathbf{2}$. $\triangle^H_{C^{[*]}}(A^{[*]}) = \{\bar{z} :\; z \in \triangle_C^H(A)\}$;

$\mathbf{D}_\mathbf{C}^\mathbf{H}\mathbf{3}$. $\triangle_C^H(A) = \triangle_A^H(C)$, isto é, são simétricos os papéis das matrizes A e C.

Repetindo o racíocinio da demonstração do Teorema 1.3.6, prova-se que a fronteira de $\triangle_C^H(A)$ é uma união finita de arcos algébricos, quando $A, C \in M_n$ são matrizes diagonais. Recorrendo ao Teorema de Tarski, pode ainda estender-se este resultado a matrizes arbitrárias $A, C \in M_n$.

Teorema 1.4.1 *Sejam $A, C \in M_n$. A fronteira de $\triangle_C^H(A)$ é uma união finita de arcos algébricos e, portanto, uma curva de classe $\mathbb{C}^\infty$, excepto possivelmente num número finito de pontos.*

Como já observámos, pode reduzir-se o estudo de $W_C^H(A)$ a $W_{C_R}^J(A_R)$, para certas matrizes A_R e C_R, sendo $J = I_r \oplus -I_{n-r}$ e r é o número de valores próprios positivos de H, contando multiplicidades. Análogo raciocínio permite simplificar o estudo de $\triangle_C^H(A)$. De facto, a partir da *lei de inércia de Sylvester* [80], verifica-se a proveitosa relação:

$$\triangle_C^H(A) \;=\; \triangle_{C_R}^J(A_R), \tag{1.62}$$

onde R é uma matriz não-singular tal que $R^*HR = J$ é a matriz de inércia de H, $A_R = R^{-1}AR$ e $C_R = R^{-1}CR$.

Sejam A e C matrizes de valores próprios, respectivamente, α_i e c_i, $i = 1, \ldots, n$. Definem-se os pontos-σ de $\triangle_C^J(A)$ por

$$z_\sigma \;=\; \prod_{i=1}^{r} \left(c_i + \alpha_{\sigma_1(i)}\right) \prod_{i=r+1}^{n} \left(c_i + \alpha_{\sigma_2(i)}\right),$$

onde $\sigma = \sigma_1\sigma_2 \in S_n$, para σ_1 e σ_2 permutações de $1, \ldots, r$ e $r+1, \ldots, n$, respectivamente. Como ilustraremos nas subsecções seguintes, o papel dos pontos-σ de $\triangle_C^J(A)$ é tão importante como o papel dos pontos-σ dos conjuntos $\triangle_C(A)$, $W_C(A)$ e $W_C^J(A)$.

Observa-se, ainda, que $\triangle_C^J(A)$ é um conjunto conexo, quaisquer que sejam $A, C \in M_n$, visto que $\triangle_C^J(A)$ é o contradomínio de uma aplicacão contínua do conjunto conexo $U_{r,n-r}$ em $\mathbb{C}$.

1.4.1 Formas Especiais do Contradomínio Determinantal-C, J

A analogia que existe entre $W_C(A)$ e $\triangle_C(A)$ e que se manifesta pela partilha de resultados similares, como a elipticidade da fronteira no caso 2×2 (excepto os focos distintos) e a identificação de pontos angulosos como pontos-σ [9, 49] estende-se aos conjuntos $W_C^J(A)$ e $\triangle_C^J(A)$. Terminamos este primeiro capítulo com alguns resultados relativos ao contradomínio determinantal-C, J, versando este tema, paralelos aos antes apresentados para o contradomínio tracial-C, J.

O caso 2×2 para $\triangle_C^J(A)$ e C diagonal

Mesmo que C seja diagonal, o estudo da geometria de $\triangle_C^J(A)$, como a de $W_C^J(A)$, revela-se tanto mais complicada quanto maior for a dimensão das matrizes em causa. O caso bidimensional é excepcionalmente simples.

Dadas $C = \operatorname{diag}(c_1, c_2) \in M_2$ e $A \in M_2$ de valores próprios α_1 e α_2, o Teorema do Contradomínio Elíptico para $\triangle_C(A)$ [12] estabelece que este conjunto é um disco elíptico (possivelmente degenerado), cujos focos são os pontos-σ $z_{\mathrm{id}} = (c_1 + \alpha_1)(c_2 + \alpha_2)$ e $z_{(12)} = (c_1 + \alpha_2)(c_2 + \alpha_1)$, e cujos eixos maior e menor têm comprimento

$$|c_1 - c_2|\sqrt{\operatorname{Tr}(A^*A) - 2\operatorname{Re}(\bar{\alpha}_1\alpha_2)} \qquad \text{e} \qquad |c_1 - c_2|\sqrt{\operatorname{Tr}(A^*A) - |\alpha_1|^2 - |\alpha_2|^2},$$

respectivamente.

Analisamos, agora, o caso bidimensional para o contradomínio determinantal-C, J, quando $J = \operatorname{diag}(1, -1)$ e C é diagonal.

Teorema 1.4.2 *Sejam $J = \operatorname{diag}(1, -1)$, $C = \operatorname{diag}(c_1, c_2) \in M_2$ e $A = [a_{ij}] \in M_2$, de valores próprios α_1 e α_2, tais que $\alpha = \det(A + C)$, $\alpha' = \det(A + \operatorname{diag}(c_2, c_1))$. Sejam*

$$\begin{aligned} M &= (c_1 - c_2)^2 \left(|\alpha_1|^2 + |\alpha_2|^2 - \operatorname{Tr}\big(A^{[*]}A\big)\right), \\ N &= (c_1 - c_2)^2 \left(\operatorname{Tr}\big(A^{[*]}A\big) - 2\operatorname{Re}(\bar{\alpha}_1\alpha_2)\right). \end{aligned}$$

Consideram-se ainda os pontos-σ $z_{\mathrm{id}} = (c_1 + \alpha_1)(c_2 + \alpha_2)$ e $z_{(12)} = (c_1 + \alpha_2)(c_2 + \alpha_1)$.

a) *Se $M > 0$ e $N > 0$, então $\triangle_C^J(A)$ é limitado por um dos ramos da hipérbole de focos z_{id} e $z_{(12)}$, cujos eixos transverso e não-transverso têm comprimento $\sqrt{N}$ e $\sqrt{M}$, respectivamente.*

b) *Se $M > 0$ e $N = 0$, então $\triangle_C^J(A)$ é:*

- **i.** *a recta l, se $|a_{12}| = |a_{21}|$,*
- **ii.** *um semi-plano aberto definido pela recta l, se $|a_{12}| \neq |a_{21}|$,*

em que l é a recta que passa por α e é perpendicular ao segmento que une z_{id} a $z_{(12)}$.

c) *Se $M > 0$ e $N < 0$, então $\triangle_C^J(A)$ é todo o plano complexo.*

d) *Se $M = 0$ e $N > 0$, então $\triangle_C^J(A)$ é uma semi-recta fechada, sobre a recta definida por z_{id} e $z_{(12)}$, com um dos pontos-σ por extremo.*

e) *Se $M = N = 0$, então $\triangle_C^J(A)$ é:*

- **i.** *o conjunto singular $\{\alpha\}$, se $a_{11} = a_{22}$ ou $c_1 = c_2$;*
- **ii.** *uma semi-recta aberta, sobre a recta definida pelos pontos α e α', de extremo α, se $a_{11} \neq a_{22}$ e $c_1 \neq c_2$.*

Demonstração: Facilmente se verifica o seguinte desenvolvimento para o determinante da soma de duas matrizes 2×2:

$$\det\left(C + U^{[*]}AU\right) \;=\; \det(C) \;+\; \det\left(U^{[*]}AU\right) \;+\; \mathrm{Tr}\left(C'U^{[*]}AU\right),$$

em que $C' = \mathrm{diag}\,(c_2, c_1)$, para qualquer matriz pseudo-unitária $U \in U_{1,1}$. Usando este desenvolvimento, tem-se

$$\triangle_C^J(A) \;=\; c_1c_2 \;+\; \alpha_1\alpha_2 \;+\; W_{C'}^J(A),$$

ou seja, o contradomínio determinantal-C, J resulta do contradomínio tracial-C', J por uma translação. Pelo Teorema 1.3.2, mediante cálculos simples, tem-se o pretendido. ■

Condições para $\triangle_C^J(A)$ ser um Conjunto Singular

Concluído o estudo do caso bidimensional, seja $J = I_r \oplus -I_{n-r}$, $0 \leq r \leq n$, e $C \in M_n$ uma matriz diagonal. Provar um lema auxiliar do tipo do Lema 1.3.1 é o nosso próximo passo, com vista a caracterizar em que condições o contradomínio determinantal-C, J de uma matriz diagonal $A \in M_n$ toma a forma de conjunto singular ou de subconjunto da recta real.

Lema 1.4.1 *Sejam* $A = \mathrm{diag}\,(a_1, \ldots, a_n) \in M_n$ *e* $C = \mathrm{diag}\,(c_1, \ldots, c_n) \in M_n$. *Então* $\triangle_C^J(A)$ *contém o conjunto*

$$\triangle_{C_{kl}}^{J_{kl}}(A_{kl}) \prod_{i \neq k,l} (a_i + c_i)\,, \tag{1.63}$$

onde $A_{kl} = A[kl]$, $C_{kl} = C[kl]$ *e* $J_{kl} = J[kl]$, $1 \leq k < l \leq n$.

Demonstração: Qualquer elemento z do conjunto (1.63) é da forma

$$z \;=\; \det\left(C_{kl} \;+\; M^{-1}A_{kl}M\right) \prod_{i \neq k,l} (a_i + c_i)\,,$$

em que $M \in M_2$ satisfaz $M^*J_{kl}M = J_{kl}$. A matriz U_M que se obtém da matriz identidade I_n substituindo as linhas e colunas k e l, pela linhas e colunas 1 e 2 da matriz M, respectivamente, satisfaz a condição $U_M^*JU_M = J$. Dadas A_2 e C_2 as submatrizes que resultam de A e C, respectivamente, eliminando as linhas e colunas k e l, verifica-se que

$$z \;=\; \det\left(C_{kl} + M^{-1}A_{kl}M\right)\det(A_2 + C_2) \;=\; \det\left(C + U_M^{[*]}AU_M\right),$$

por isso, $z \in \triangle_C^J(A)$. ■

Se pelo menos uma das matrizes A ou C for escalar, então $\triangle_C^J(A) = \{\det(A + C)\}$.

Teorema 1.4.3 *Sejam $A, C \in M_n$ matrizes diagonais e C não escalar, tais que os pontos-σ de $\triangle_C^J(A)$ são não nulos. O conjunto $\triangle_C^J(A)$ é singular se e só se A é uma matriz escalar.*

Demonstração: A implicação ($\Leftarrow$) é óbvia.

($\Rightarrow$) Suponhamos que $A = \operatorname{diag}(a_1, \ldots, a_n)$ é não escalar. Existem pelo menos duas entradas diagonais distintas a_k e a_l, $1 \leq k < l \leq n$ (com $k \leq r < l$, se $0 < r < n$). Seja $C_\sigma = P_\sigma C P_\sigma^T$, em que $C = \operatorname{diag}(c_1, \ldots, c_n)$ e P_σ é a matriz de permutação associada a $\sigma = \sigma_1\sigma_2 \in S_n$, com σ_1 e σ_2 permutações de $1, \ldots, r$ e $r+1, \ldots, n$, respectivamente, de tal modo que $C'_{kl} = C_\sigma[kl]$ é não escalar. Pela propriedade $D_{\mathrm{C}}^{\mathrm{H}}1$, tem-se $\triangle_C^J(A) = \triangle_{C_\sigma}^J(A)$. Pelo Lema 1.4.1, $\triangle_C^J(A)$ contém o conjunto

$$\Pi_{kl} \;=\; \triangle_{C'_{kl}}^{J_{kl}}(A_{kl}) \prod_{i \neq k,l} \left(a_i + c_{\sigma(i)}\right).$$

onde $A_{kl} = A[kl]$ e $J_{kl} = J[kl]$. Pelo Teorema do Contradomínio Elíptico relativo a $\triangle_C(A)$ e pelo Teorema 1.4.2, o subconjunto Π_{kl} de $\triangle_C^J(A)$ é um disco elíptico, se $j_k = j_l$, ou é um ramo de uma hipérbole com interior, se $j_k \neq j_l$. Como A_{kl}, C'_{kl} são matrizes não escalares e os pontos-σ de $\triangle_C^J(A)$ são não nulos, em ambos os casos, Π_{kl} pode degenerar, mas nunca num só ponto e, portanto, $\triangle_C^J(A)$ não é um conjunto singular. ■

Condições para $\triangle_C^J(A)$ ser um Subconjunto da Recta Real

Argumentos análogos aos da demonstração do Teorema 1.4.3, permitem provar o seguinte resultado.

Teorema 1.4.4 *Sejam $A, C \in M_n$ matrizes diagonais e $C \in M_n(\mathbb{R})$ não escalar, tais que os pontos-σ de $\triangle_C^J(A)$ são não nulos. Então $\triangle_C^J(A)$ é um subconjunto do eixo real se e só se $A \in M_n(\mathbb{R})$.*

Demonstração: ($\Leftarrow$) Todo o $z \in \triangle_C^J(A)$ é da forma $z = \det\left(C + U^{[*]}AU\right)$, para uma certa matriz $U \in \mathcal{U}_{r,n-r}$, e satisfaz

$$\bar{z} \;=\; \det\left(C + U^{[*]}AU\right)^* \;=\; \det\left(C^* + U^*A^*JUJ\right) \;=\; \det\left(C^{[*]} + U^{[*]}A^{[*]}U\right).$$

Como $A, C \in M_n(\mathbb{R})$ são diagonais, $A^{[*]} = A$, $C^{[*]} = C$ e $\bar{z} = z$, donde $\triangle_C^J(A) \subseteq \mathbb{R}$.

($\Rightarrow$) Suponhamos que $A = \operatorname{diag}(a_1, \ldots, a_n)$ tem pelo menos uma entrada não real. Então A contém uma submatriz principal $A_{kl} = \operatorname{diag}(a_k, a_l)$ que não é Hermítica-J_{kl}, onde $J_{kl} = \mathrm{diag}\,(j_k, j_l)$, $1 \leq k < l \leq n$ (com $k \leq r < l$, se $0 < r < n$). Como na demonstração do Teorema 1.4.3 e atendendo, de novo, ao Lema 1.4.1, $\triangle_C^J(A)$ contém

$$\Pi_{kl} \;=\; \triangle_{C'_{kl}}^{J_{kl}}(A_{kl}) \prod_{i \neq k,l} \left(a_i + c_{\sigma(i)}\right).$$

onde $\sigma = \sigma_1\sigma_2$, com σ_1 e σ_2 permutações de $1, \ldots, r$ e $r+1, \ldots, n$, respectivamente, tal que $C'_{kl} = \mathrm{diag}\left(c_{\sigma(k)}, c_{\sigma(l)}\right)$ é não escalar. Como A_{kl} não é Hermítica-J_{kl} e os pontos-σ de $\triangle_C^J(A)$ não se anulam, similarmente à demonstração do Teorema 1.4.3, verifica-se que o subconjunto Π_{kl} de $\triangle_C^J(A)$ é um disco elíptico ou um ramo de uma hipérbole com interior, que nunca degenera num subconjunto do eixo real. ■

Apesar das propriedades semelhantes partilhadas pelos conjuntos $W_C^J(A)$ e $\triangle_C^J(A)$, o nível de dificuldade dos problemas que envolvem o segundo é manifestamente superior, daí serem menos gerais as condições dos teoremas obtidos neste último caso.

1.4.2 Pontos Angulosos do Contradomínio Determinantal-C, J

Natália Bebiano [12] investigou os pontos angulosos não nulos de $\triangle_C(A)$, quando C é diagonal. Neste secção, considera-se $J = I_r \oplus -I_{n-r}$, $0 \leq r \leq n$, e obtém-se um resultado análogo para os pontos angulosos não nulos de $\triangle_C^J(A)$. Antes, referimos o seguinte lema devido a Fiedler [52].

Lema 1.4.2 *Se $X, Y \in M_n$ são tais que X é não-singular, então*

$$\det(X + t\,Y) \;=\; \det(X)\left[1 + t\,\mathrm{Tr}\left(Y X^{-1}\right)\right] + O(t^2), \tag{1.64}$$

para qualquer $t \in \mathbb{R}$ numa vizinhança da origem.

Demonstração: Facilmente se verifica (1.64), quando $X = I_n$, a partir do desenvolvimento do polinómio característico de uma matriz $A \in M_n$, isto é,

$$p_A(t) \;=\; t^n \;+\; \sum_{k=1}^{n} (-1)^k E_k(A)\, t^{n-k},$$

em que $E_k(A)$ denota a soma dos $\binom{n}{k}$ menores principais de A de ordem k. De facto, para qualquer $t \in \mathbb{R}$ numa vizinhança da origem, tem-se

$$\det(I_n + t\,A) \;=\; t^n\, p_{-A}(1/t) \;=\; 1 \;+\; t\,\mathrm{Tr}\,(A) + O(t^2),$$

atendendo a que $E_1(-A) = -\mathrm{Tr}\,(A)$. Em geral, basta tomar $A = Y X^{-1}$ e notar que $\det(X + t\,Y) = \det(X)\det(I_n + t\,Y X^{-1})$. ■

Teorema 1.4.5 *Seja $C = c_1 I_{n_1} \oplus \cdots \oplus c_p I_{n_p}$, $n_1, \ldots, n_p \in \mathbb{N}$, $n_1 + \cdots + n_p = n$, com $c_1, \ldots, c_p$ números complexos distintos dois a dois. Se $A \in M_n$ e $z = \det\big(C + U^{[*]}AU\big)$, $U \in U_{r,n-r}$, é um ponto anguloso não nulo de $\triangle_C^J(A)$, então*

$$U^{[*]}AU \;=\; A_1 \oplus \cdots \oplus A_p \quad e \quad z \;=\; \prod_{i=1}^{p} \det\left(c_i I_{n_i} + A_i\right), \qquad A_i \in M_{n_i}, \quad i = 1, \ldots, p.$$

Demonstração: Por simplicidade de notação, escreve-se $A_U = U^{[*]}AU$. Dada $S \in M_n$ Hermítica, considera-se a função diferenciável

$$f(t) \;=\; \det\left(A_U + \mathrm{e}^{itJ^{1/2}SJ^{1/2}} C\, \mathrm{e}^{-itJ^{1/2}SJ^{1/2}}\right),$$

que, pelo Lema 1.4.2, admite o seguinte desenvolvimento

$$f(t) \;=\; \det\left(C + A_U\right)\left[1 + i\,t\,\mathrm{Tr}\left((C + A_U)^{-1}\left(J^{1/2}SJ^{1/2}C - CJ^{1/2}SJ^{1/2}\right)\right)\right] + O(t^2),$$

para $t \in \mathbb{R}$ numa vizinhança de zero. Como $\det\left(C + A_U\right) \neq 0$ é um ponto anguloso de $\triangle_C^J(A)$, a derivada de $f(t)$ em relação a t na origem anula-se, logo

$$\mathrm{Tr}\left(SJ^{1/2}\left[C, (C + A_U)^{-1}\right] J^{1/2}\right) \;=\; 0. \tag{1.65}$$

Sendo S arbitrária, de (1.65), conclui-se que $\left[C, (C + A_U)^{-1}\right] = 0$, o que implica que $[C, A_U] = 0$. Logo A_U e C são simultaneamente diagonalizáveis por uma transformação de semelhança unitária. Repetindo o racíocinio do final da demonstração do Teorema 1.3.7, tiram-se as conclusões pretendidas. ∎

Como consequência do Teorema 1.4.5, quando $p = n$, temos o seguinte resultado "dual" do Corolário 1.3.5, com excepção do ponto zero.

Corolário 1.4.1 *Seja $C = \mathrm{diag}\left(c_1, \ldots, c_n\right) \in M_n$, com $c_1, \ldots, c_n$ todos distintos. Se $A \in M_n$ tem valores próprios $\alpha_1, \ldots, \alpha_n$ e $z \in \triangle_C^J(A)$ é um ponto anguloso não nulo de $\triangle_C^J(A)$, então z é um ponto-σ do tipo*

$$z_\sigma \;=\; \prod_{i=1}^{r}\left(c_i + \alpha_{\sigma_1(i)}\right) \prod_{i=r+1}^{n}\left(c_i + \alpha_{\sigma_2(i)}\right), \tag{1.66}$$

com $\sigma = \sigma_1\sigma_2 \in S_n$, para σ_1 e σ_2 permutações de $1, \ldots, r$ e $r+1, \ldots, n$, respectivamente.

Capítulo 2

Contradomínios Numéricos em Física

Symmetry, as wide or as narrow as you may define it, is one idea by which man through the ages has tried to comprehend order, beauty and perfection.

Herman Weyl, in *Introductory Statistichal Mechanics*

The miracle of the appropriateness of the language of mathematics for the formulation of the laws of physics is a wonderful gift which we neither understand nor deserve.

Eugene P. Wigner (Prémio Nobel da Física)

Em Mecânica Quântica, descrevem-se os estados de uma partícula por vectores de um espaço de Hilbert, dito o *espaço de estados*. Perante sistemas físicos compostos por várias partículas idênticas, definir operadores que criam ou destroem partículas em estados individuais específicos revela-se especialmente útil. Outros operadores de interesse físico podem exprimir-se em termos destes operadores de criação e destruição. Neste segundo capítulo, investigamos o contradomínio numérico de certos operadores deste tipo, alguns dos quais ilimitados, mas que admitem representações matriciais tridiagonais infinitas bem estruturadas. Antes de mais, introduzimos alguns conceitos preliminares aos espaços de estados adequados a este estudo.

2.1 Preliminares de Classes Simétricas de Tensores

Seja $\mathcal{H}$ um espaço de Hilbert complexo n-dimensional com produto interno $\langle \cdot, \cdot \rangle$ e $m \in \mathbb{N}$. Denota-se por $\otimes^m \mathcal{H}$ o m-ésimo produto tensorial de $\mathcal{H}$ e por $x_1 \otimes \cdots \otimes x_m$ o produto tensorial dos vectores $x_1, \ldots, x_m \in \mathcal{H}$.

Seja G um subgrupo de S_m, o grupo simétrico de grau m, de ordem $|G|$ e seja id a permutação identidade em S_m. Uma função $\chi : G \to \mathbb{C}$ satisfazendo $\chi(\mathrm{id}) = 1$ e $\chi(\sigma_1\sigma_2) = \chi(\sigma_1)\,\chi(\sigma_2)$, quaisquer que sejam $\sigma_1, \sigma_2 \in G$, diz-se um *caracter complexo de grau um* em G. O único operador linear $S_\chi^G : \otimes^m\mathcal{H} \to \otimes^m\mathcal{H}$ que satisfaz

$$S_\chi^G(x_1 \otimes \cdots \otimes x_m) \;=\; |G|^{-1/2} \sum_{\sigma\in G} \chi(\sigma)\; x_{\sigma^{-1}(1)} \otimes \cdots \otimes x_{\sigma^{-1}(m)} \tag{2.1}$$

designa-se o *simetrizador* determinado por G e χ. A imagem de $\otimes^m\mathcal{H}$ pelo simetrizador S_χ^G denomina-se *classe simétrica de tensores* sobre $\mathcal{H}$ associada a G e χ, e denota-se por $\mathcal{H}_\chi^m(G)$. Os elementos da forma (2.1) são os *tensores decomponíveis* de $\mathcal{H}_\chi^m(G)$.

O estudo de classes simétricas de tensores tem motivações de várias áreas da matemática pura e aplicada: teoria das matrizes, teoria dos operadores, teoria combinatória, representação de grupos, geometria diferencial, geometria algébrica, equações com derivadas parciais, mecânica quântica e outras [101].

Com vista a munir o espaço $\mathcal{H}_\chi^m(G)$ de um produto interno, define-se a *função generalizada de matrizes* $d_\chi^G\colon M_m \to \mathbb{C}$ associada a G e χ por

$$d_\chi^G(X) \;=\; \sum_{\sigma\in G} \chi(\sigma)\, x_{1\sigma(1)} \cdots x_{m\sigma(m)}, \qquad X = [x_{ij}] \in M_m.$$

Induzido pelo produto interno em $\mathcal{H}$, dota-se $\mathcal{H}_\chi^m(G)$ de um produto interno definido por

$$\langle S_\chi^G(x_1 \otimes \cdots \otimes x_m)\,, S_\chi^G(y_1 \otimes \cdots \otimes y_m)\,\rangle_m \;=\; d_\chi^G[\langle x_i, y_j\rangle],$$

para $x_1 \otimes \cdots \otimes x_m$ e $y_1 \otimes \cdots \otimes y_m$ vectores decomponíveis em $\otimes^m\mathcal{H}$. Identificando o espaço de Hilbert $\mathcal{H}$ com $\mathbb{C}^n$ e cada tensor de $\mathcal{H}_\chi^m(G)$ do tipo (2.1) com uma matriz X de colunas $x_1, \ldots, x_m$, a norma $\|\cdot\|$ induzida pelo produto interno em $\mathcal{H}_\chi^m(G)$ é tal que $\|S_\chi^G(x_1 \otimes \cdots \otimes x_m)\|^2 = d_\chi^G(X^*X)$.

Introduz-se seguidamente alguma notação que será útil na construção de uma base (ortonormada) do espaço $\mathcal{H}_\chi^m(G)$ a partir de uma base (ortonormada) de $\mathcal{H}$. (Para um tratamento aprofundado sobre este assunto, consultar [101].)

Considera-se o conjunto $\Gamma_{m,n}$ das sucessões de comprimento m de números inteiros de 1 a n. O número de ocorrências do inteiro i na sucessão $\alpha \in \Gamma_{m,n}$ denomina-se o *número de ocupação de nível* i de α e denota-se por $k_i(\alpha)$.

Duas sucessões α e β em $\Gamma_{m,n}$ dizem-se *equivalentes módulo* G, o que se denota por $\alpha \overset{G}{\sim} \beta$, se existir uma permutação $\sigma \in G$ tal que $\beta = \alpha\sigma$. Por exemplo, a relação de equivalência módulo S_m pode ser explicitada, à custa dos números de ocupação das sucessões em causa, da seguinte forma: $\alpha \overset{S_m}{\sim} \beta$ se e só se $k_i(\alpha) = k_i(\beta)$, $i = 1, \ldots, n$.

Obviamente, a relação de equivalência $\overset{G}{\sim}$ particiona $\Gamma_{m,n}$ em classes de equivalência. Considera-se um sistema de representantes ∇ destas classes de equivalência, em que se escolhe cada sucessão de ∇ como a primeira, na ordem lexicográfica, da sua classe de equivalência. Na relação módulo S_m, tal significa escolher, em cada classe de equivalência distinta, a sucessão de componentes $\alpha_1, \ldots, \alpha_m$ que satisfaz $1 \leq \alpha_1 \leq \cdots \leq \alpha_m \leq n$; portanto, ∇ é o conjunto $G_{m,n}$ das sucessões não decrescentes em $\Gamma_{m,n}$.

Chama-se *estabilizador* de $\alpha \in \Gamma_{m,n}$ ao subgrupo $G_\alpha = \{\sigma \in G : \alpha\sigma = \alpha\}$ de G. Define-se $\bar{\nabla}$ como o subconjunto de ∇ das sucessões cujo estabilizador está contido no núcleo do caracter χ, ou seja, das sucessões $\alpha \in \nabla$ tais que $|G_\alpha| = \sum_{\sigma \in G_\alpha} \chi(\sigma)$.

Dada uma base $\{e_1, \ldots, e_n\}$ de $\mathcal{H}$, escreve-se $e_\alpha^\otimes = e_{\alpha_1} \otimes \cdots \otimes e_{\alpha_m}$, para cada sucessão $\alpha \in \Gamma_{m,n}$ de componentes $\alpha_1, \ldots, \alpha_m$. O conjunto $\left\{S_\chi^G(e_\alpha^\otimes) : \alpha \in \bar{\nabla}\right\}$ constitui uma base da classe simétrica de tensores $\mathcal{H}_\chi^m(G)$. Se a base $\{e_1, \ldots, e_n\}$ de $\mathcal{H}$ for ortonormada, então $\left\{|G_\alpha|^{-1/2} S_\chi^G(e_\alpha^\otimes) : \alpha \in \bar{\nabla}\right\}$ é uma base ortonormada de $\mathcal{H}_\chi^m(G)$.

O produto tensorial $\otimes^m \mathcal{H}$ é uma classe simétrica de tensores associada ao subgrupo trivial $G = \{\mathrm{id}\}$. Além disso, se $\{e_1, \ldots, e_n\}$ é uma base (ortonormada) de $\mathcal{H}$, então $\nabla = \bar{\nabla} = \Gamma_{m,n}$ e $\{e_\alpha^\otimes : \alpha \in \Gamma_{m,n}\}$ constitui uma base (ortonormada) de $\otimes^m \mathcal{H}$.

Apresentam-se, de seguida, dois outros exemplos de classes simétricas de tensores, ambas associadas ao grupo simétrico de grau m, de interesse especial pelas suas aplicações a variadas áreas da matemática e da física.

Espaço Simétrico

A classe simétrica de tensores sobre $\mathcal{H}$ associada ao grupo simétrico S_m e ao caracter principal 1 denomina-se o *m-ésimo espaço (completamente) simétrico* sobre $\mathcal{H}$ e denota-se por $\mathcal{H}_{(m)}$, $m \in \mathbb{N}$. Escreve-se

$$x_1 * \cdots * x_m \; = \; (m!)^{-1/2} \sum_{\sigma \in S_m} x_{\sigma^{-1}(1)} \otimes \cdots \otimes x_{\sigma^{-1}(m)}$$

para denotar um elemento decomponível de $\mathcal{H}_{(m)}$, dito o *produto (tensorial) simétrico* dos vectores $x_1, \ldots, x_m \in \mathcal{H}$. Neste caso, como a função generalizada de matrizes $d_1^{S_m}$ é dada pelo permanente, define-se um produto interno por

$$\langle x_1 * \cdots * x_m, y_1 * \cdots * y_m \rangle_m \; = \; \operatorname{per}\left[\langle x_i, y_j \rangle\right],$$

para $x_1 * \cdots * x_m$ e $y_1 * \cdots * y_m$ vectores decomponíveis de $\mathcal{H}_{(m)}$.

Perante a relação módulo S_m, o sistema de representantes ∇ das classes de equivalência é o conjunto $G_{m,n}$ e, como $\chi = 1$, tem-se $\bar{\nabla} = \nabla$. Dada uma base $\{e_1, \ldots, e_n\}$ de $\mathcal{H}$, escreve-se $e^*_\alpha = e_{\alpha_1} * \cdots * e_{\alpha_m}$, para cada sucessão $\alpha \in G_{m,n}$ de componentes $\alpha_1, \ldots, \alpha_m$. O conjunto

$$\{e^*_\alpha : \ \alpha \in G_{m,n}\} \tag{2.2}$$

constitui uma base de $\mathcal{H}_{(m)}$. Se a base $\{e_1, \ldots, e_n\}$ de $\mathcal{H}$ for ortonormada, então os vectores de (2.2) previamente divididos por $\sqrt{k_1(\alpha)\,! \cdots k_n(\alpha)\,!}$ constituem uma base ortonormada de $\mathcal{H}_{(m)}$. A dimensão do espaço $\mathcal{H}_{(m)}$ é $\binom{m+n-1}{n}$.

É adequado fazer-se a seguinte interpretação física. Suponhamos que existem m partículas num sistema físico, obedecendo à estatística de Bose-Einstein, cujos vectores de estado normalizados $x_1, \ldots, x_n$ formam um conjunto ortonormado. Supondo que existem exactamente k_j partículas no estado x_j, $j = 1, \ldots, n$, então representa-se o seu estado misto pelo tensor $x^*_\alpha = x_{\alpha_1} * \cdots * x_{\alpha_m}$, onde a sucessão $\alpha \in G_{m,n}$ de componentes $\alpha_1, \ldots, \alpha_m$ tem números de ocupação $k_1, \ldots, k_n$ [18].

Por convenção, $\mathcal{H}_{(0)} = \mathbb{C}$. A soma directa dos m-ésimos espaços simétricos sobre $\mathcal{H}$, $\bigoplus_{m=0}^{+\infty} \mathcal{H}_{(m)}$, designa-se a *álgebra simétrica* sobre $\mathcal{H}$, e denota-se por $\Gamma^*(\mathcal{H})$.

Espaço de Grassmann

O caracter $\varepsilon : S_m \to \mathbb{C}$, onde $\varepsilon(\sigma)$ é o sinal da permutação $\sigma \in S_m$, diz-se o *caracter alternante*. Denota-se por $\bigwedge^m \mathcal{H}$, $m \in \mathbb{N}$, a classe simétrica de tensores sobre $\mathcal{H}$ associada a S_m e ao caracter alternante. A $\bigwedge^m \mathcal{H}$ dá-se o nome de *m-ésimo espaço de Grassmann* sobre $\mathcal{H}$, também conhecido por *m-ésimo espaço exterior* de $\mathcal{H}$ ou *m-ésimo espaço anti--simétrico* sobre $\mathcal{H}$. Escreve-se

$$x_1 \wedge \cdots \wedge x_m \ = \ (m!)^{-1/2} \sum_{\sigma \in S_m} \varepsilon(\sigma)\ x_{\sigma^{-1}(1)} \otimes \cdots \otimes x_{\sigma^{-1}(m)}$$

para denotar um seu elemento decomponível, dito o *produto (tensorial) anti-simétrico* ou o *produto exterior* dos vectores $x_1, \ldots, x_m \in \mathcal{H}$. Neste caso, a função generalizada $d^{S_m}_\varepsilon$ é o determinante e define-se um produto interno por

$$\langle x_1 \wedge \cdots \wedge x_m, y_1 \wedge \cdots \wedge y_m \rangle_m \ = \ \det\,[\langle x_i, y_j \rangle],$$

para $x_1 \wedge \cdots \wedge x_m$ e $y_1 \wedge \cdots \wedge y_m$ vectores decomponíveis de $\bigwedge^m \mathcal{H}$.

Perante a relação módulo S_m, tem-se $\nabla = G_{m,n}$. Sendo $\chi = \varepsilon$, verifica-se que $\bar{\nabla}$ é o conjunto $Q_{m,n}$ das sucessões estritamente crescentes em $\Gamma_{m,n}$. Se $m > n$, então $Q_{m,n} = \emptyset$, razão pela qual apenas interessa considerar $\bigwedge^m \mathcal{H}$, quando $m \le n$. Dada uma base (ortonormada) $\{e_1, \ldots, e_n\}$ de $\mathcal{H}$, escreve-se $e^\wedge_\alpha = e_{\alpha_1} \wedge \cdots \wedge e_{\alpha_m}$, para cada sucessão $\alpha \in Q_{m,n}$ de componentes $\alpha_1, \ldots, \alpha_m$. Ora, o estabilizador de $\alpha \in Q_{m,n}$ tem ordem um e

$$\{e^\wedge_\alpha : \ \alpha \in Q_{m,n}\} \tag{2.3}$$

constitui uma base (ortonormada) de $\bigwedge^m \mathcal{H}$. Logo, $\bigwedge^m \mathcal{H}$ é de dimensão $\binom{n}{m}$.

Se $m = 0$, convenciona-se que $\bigwedge^0 \mathcal{H} = \mathbb{C}$. A soma directa dos m-ésimos espaços de Grassmann sobre $\mathcal{H}$, $\bigoplus_{m=0}^n \bigwedge^m \mathcal{H}$, diz-se a *álgebra de Grassmann* sobre $\mathcal{H}$ e denota-se por $\Gamma^\wedge(\mathcal{H})$. Esta álgebra é de dimensão finita 2^n.

2.2 Operadores de Criação e de Destruição

Em Física Quântica, podem exprimir-se operadores lineares a actuar em espaços de Hilbert de sistemas de muitos-corpos a partir de operadores de criação e destruição [29, 33, 110]. Na Natureza, observam-se estados completamente simétricos e completamente anti-simétricos; as partículas que ocorrem nestes estados chamam-se *bosões* e *fermiões*, respectivamente. As primeiras caracterizam-se por um seu número arbitrário poder ocupar o mesmo estado quântico e as segundas pela impossibilidade de dois fermiões do sistema ocuparem o mesmo estado. Se $\mathcal{H}$ for o espaço de estados de uma só partícula, o m-ésimo espaço simétrico sobre $\mathcal{H}$ e o m-ésimo espaço de Grassmann sobre $\mathcal{H}$ são os espaços de estados adequados para descrever sistemas com m bosões e m fermiões, respectivamente. Importa, assim, considerar separadamente operadores de criação e destruição de bosões e de fermiões. Salientaremos as semelhanças que entre eles ocorram.

2.2.1 Operadores de Criação e de Destruição de Bosões

Seja $\{e_1, \ldots, e_n\}$ uma base ortonormada de $\mathcal{H}$. Introduz-se o conceito de *operador de criação bosónico* $f_i : \mathcal{H}_{(m)} \to \mathcal{H}_{(m+1)}$, associado ao vector e_i, $i = 1, \ldots, n$, como o operador linear definido por

$$f_i(x_1 * \cdots * x_m) \;=\; e_i * x_1 * \cdots * x_m,$$

para cada vector decomponível $x_1 * \cdots * x_m \in \mathcal{H}_{(m)}$. Chama-se *operador de destruição bosónico* ao operador adjunto g_i do operador de criação bosónico f_i, $i = 1, \ldots, n$.

Proposição 2.2.1 *O operador adjunto do operador de criação $f_i : \mathcal{H}_{(m)} \to \mathcal{H}_{(m+1)}$ é o operador linear $g_i : \mathcal{H}_{(m+1)} \to \mathcal{H}_{(m)}$ definido por*

$$g_i(x_1 * \cdots * x_{m+1}) \;=\; \sum_{k=1}^{m+1} \langle e_i, x_k \rangle \, x_1 * \cdots * x_{k-1} * x_{k+1} * \cdots * x_{m+1},$$

para cada vector decomponível $x_1 * \cdots * x_{m+1} \in \mathcal{H}_{(m+1)}$.

Demonstração: O adjunto de $f_i : \mathcal{H}_{(m)} \to \mathcal{H}_{(m+1)}$ é o operador linear $g_i : \mathcal{H}_{(m+1)} \to \mathcal{H}_{(m)}$ que satisfaz

$$\langle f_i(x_1 * \cdots * x_m)\, ,\, y_1 * \cdots * y_{m+1} \rangle_{m+1} \;=\; \langle\, x_1 * \cdots * x_m\, ,\, g_i(y_1 * \cdots * y_{m+1}) \,\rangle_m, \qquad (2.4)$$

para cada $x_1 * \cdots * x_m \in \mathcal{H}_{(m)}$ e $y_1 * \cdots * y_{m+1} \in \mathcal{H}_{(m+1)}$. Sendo o produto interno do primeiro membro de (2.4) dado por

$$\operatorname{per}\begin{bmatrix} \langle e_i, y_1\rangle & \cdots & \langle e_i, y_k\rangle & \cdots & \langle e_i, y_{m+1}\rangle \\ \langle x_1, y_1\rangle & \cdots & \langle x_1, y_k\rangle & \cdots & \langle x_i, y_{m+1}\rangle \\ \vdots & & \vdots & & \vdots \\ \langle x_m, y_1\rangle & \cdots & \langle x_m, y_k\rangle & \cdots & \langle x_m, y_{m+1}\rangle \end{bmatrix}$$

e atendendo ao desenvolvimento de Laplace do permanente segundo a primeira linha,

$$\sum_{k=1}^{m+1} \langle e_i, y_k\rangle\, \langle x_1 * \cdots * x_m\, ,\, y_1 * \cdots * y_{k-1} * y_{k+1} * \cdots * y_{m+1}\rangle_m,$$

obtém-se a acção do operador g_i sobre os vectores decomponíveis $y_1 * \cdots * y_{m+1} \in \mathcal{H}_{(m+1)}$:

$$g_i(y_1 * \cdots * y_{m+1}) \;=\; \sum_{k=1}^{m+1} \langle e_i, y_k\rangle\, y_1 * \cdots * y_{k-1} * y_{k+1} * \cdots * y_{m+1}. \qquad \blacksquare$$

Denotando por e_i^k o produto simétrico $e_i * \cdots * e_i$ com k factores, é evidente que

$$f_i(e_i^m) \;=\; e_i^{m+1} \qquad \text{e} \qquad g_i(e_i^{m+1}) \;=\; (m+1)\, e_i^m.$$

Apresentam-se, de imediato, algumas propriedades satisfeitas pelos operadores de criação e destruição de bosões, conhecidas por *relações de comutação canónicas*, em que $[f, g] = fg - gf$ denota, como usualmente, o *comutador* dos operadores f e g.

Proposição 2.2.2 *Os operadores de criação f_i e de destruição g_i de bosões satisfazem as seguintes relações de comutação canónicas:*

$$[f_i, f_j] \;=\; [g_i, g_j] \;=\; 0 \qquad e \qquad [g_i, f_j] \;=\; \delta_{ij}, \qquad i, j = 1, \ldots, n.$$

Demonstração: A título exemplificativo, prova-se a última propriedade, bastando verificá-la para cada vector decomponível $x_1 * \cdots * x_m \in \mathcal{H}_{(m)}$. Dados $i, j = 1, \ldots, n$, tem-se

$$\begin{aligned} [g_i, f_j]\, x_1 * \cdots * x_m \;&=\; g_i(e_j * x_1 * \cdots * x_m) \;-\; \sum_{k=1}^{m} \langle e_i, x_k\rangle\, e_j * x_{(k)}^* \\ &=\; \langle e_i, e_j\rangle\, x_1 * \cdots * x_m \;=\; \delta_{ij}\; x_1 * \cdots * x_m, \end{aligned} \qquad (2.5)$$

em que $x^*_{(k)}$ representa o vector $x_1 * \cdots * x_{k-1} * x_{k+1} * \cdots * x_m$ e a última igualdade em (2.5) é devida à ortonormalidade dos vectores $e_1, \ldots, e_n$. ∎

Seja $N_i : \mathcal{H}_{(m)} \to \mathcal{H}_{(m)}$ o operador linear definido por $N_i = f_i\, g_i,\ i = 1, \ldots, n$. Podem escrever-se os elementos da base (2.2) do m-ésimo espaço simétrico sobre $\mathcal{H}$ na forma $e_1^{k_1} * \cdots * e_n^{k_n}$, com $k_1, \ldots, k_n \in \mathbb{N}_0$, tais que $k_1 + \cdots + k_n = m$. É claro que $N_i : \mathcal{H}_{(m)} \to \mathcal{H}_{(m)}$ é um operador auto-adjunto e como satisfaz

$$N_i(e_1^{k_1} * \cdots * e_n^{k_n}) = k_i\, e_1^{k_1} * \cdots * e_n^{k_n}, \qquad i = 1, \ldots, n,$$

os seus valores próprios são os números inteiros não negativos não superiores a m. Por este motivo, N_i denomina-se o *operador número bosónico no estado i*. Resulta da Proposição 2.2.2 que estes operadores comutam entre si, ou seja, $[N_i, N_j] = 0,\ i,\ j = 1, \ldots, n$.

2.2.2 Operadores de Criação e de Destruição de Fermiões

O *operador de criação fermiónico*, associado ao vector e_i, $i = 1, \ldots, n$, de uma base ortonormado de $\mathcal{H}$, é o operador linear $f_i : \bigwedge^m \mathcal{H} \to \bigwedge^{m+1} \mathcal{H}$ que se define por

$$f_i(x_1 \wedge \cdots \wedge x_m) \ = \ e_i \wedge x_1 \wedge \cdots \wedge x_m,$$

para cada vector decomponível $x_1 \wedge \cdots \wedge x_m \in \bigwedge^{m+1} \mathcal{H}$. O adjunto do operador de criação fermiónico f_i diz-se o *operador de destruição fermiónico g_i*. Repetindo o raciocínio da Proposição 2.2.1, exibe-se explicitamente a acção do operador de destruição de fermiões.

Proposição 2.2.3 *O operador de destruição $g_i : \bigwedge^{m+1} \mathcal{H} \to \bigwedge^m \mathcal{H}$ é o operador linear definido por*

$$g_i(x_1 \wedge \cdots \wedge x_{m+1}) \ = \ \sum_{k=1}^{m+1} (-1)^{k+1} \langle e_i, x_k \rangle\, x_1 \wedge \cdots \wedge x_{k-1} \wedge x_{k+1} \wedge \cdots \wedge x_{m+1},$$

para cada vector decomponível $x_1 \wedge \cdots \wedge x_{m+1} \in \bigwedge^{m+1} \mathcal{H}$.

Demonstração: Se g_i é adjunto de $f_i : \bigwedge^m \mathcal{H} \to \bigwedge^{m+1} \mathcal{H}$, então $g_i : \bigwedge^{m+1} \mathcal{H} \to \bigwedge^m \mathcal{H}$ é o operador linear que satisfaz

$$\langle f_i(x_1 \wedge \cdots \wedge x_m)\, , \, y_1 \wedge \cdots \wedge y_{m+1} \rangle_{m+1} \ = \ \langle x_1 \wedge \cdots \wedge x_m\, , \, g_i(y_1 \wedge \cdots \wedge y_{m+1})\, \rangle_m, \quad (2.6)$$

quaisquer que sejam $x_1 \wedge \cdots \wedge x_m \in \bigwedge^m \mathcal{H}$ e $y_1 \wedge \cdots \wedge y_{m+1} \in \bigwedge^{m+1} \mathcal{H}$. A partir do produto interno do primeiro membro de (2.6) dado por

$$\det \begin{bmatrix} \langle e_i, y_1 \rangle & \cdots & \langle e_i, y_k \rangle & \cdots & \langle e_i, y_{m+1} \rangle \\ \langle x_1, y_1 \rangle & \cdots & \langle x_1, y_k \rangle & \cdots & \langle x_i, y_{m+1} \rangle \\ \vdots & & \vdots & & \vdots \\ \langle x_m, y_1 \rangle & \cdots & \langle x_m, y_k \rangle & \cdots & \langle x_m, y_{m+1} \rangle \end{bmatrix}$$

e do respectivo desenvolvimento do determinante segundo a primeira linha,

$$\sum_{k=1}^{m+1}(-1)^{k+1}\langle e_i, y_k\rangle\,\langle x_1\wedge\cdots\wedge x_m\,,\,y_1\wedge\cdots\wedge y_{k-1}\wedge y_{k+1}\wedge\cdots\wedge y_{m+1}\rangle_m,$$

obtém-se a acção de g_i sobre os vectores decomponíveis $y_1\wedge\cdots\wedge y_{m+1}\in\bigwedge^{m+1}\mathcal{H}$:

$$g_i(y_1\wedge\cdots\wedge y_{m+1})\;=\;\sum_{k=1}^{m+1}(-1)^{k+1}\,\langle e_i, y_k\rangle\, y_1\wedge\cdots\wedge y_{k-1}\wedge y_{k+1}\wedge\cdots\wedge y_{m+1}. \qquad\blacksquare$$

De seguida, apresentam-se as chamadas *relações de anti-comutação canónicas* satisfeitas pelos operadores de criação e destruição de fermiões, em que $\{f,g\}=fg+gf$ denota o *anti-comutador* dos operadores f e g.

Proposição 2.2.4 *Os operadores de criação f_i e de destruição g_i de fermiões satisfazem as seguintes relações de anti-comutação canónicas:*

$$\{f_i,f_j\}\;=\;\{g_i,g_j\}\;=\;0\qquad e\qquad \{g_i,f_j\}\;=\;\delta_{ij},\qquad\quad i,j=1,\ldots,n.$$

Demonstração: As primeiras relações são consequência imediata de $e_i\wedge e_j+e_j\wedge e_i=0$. Prova-se a última propriedade, sendo suficiente verificá-la para cada vector decomponível $x_1\wedge\cdots\wedge x_m\in\bigwedge^{m+1}\mathcal{H}$. Dados $i,j=1,\ldots,n$, tem-se

$$\begin{aligned} g_if_j\,(x_1\wedge\cdots\wedge x_m) &= \langle e_i,e_j\rangle\, x_1\wedge\cdots\wedge x_m\;+\;\sum_{k=1}^{m}(-1)^{k+2}\langle e_i,x_k\rangle\, e_j\wedge x^{\wedge}_{(k)}\\ &= \delta_{ij}\;x_1\wedge\cdots\wedge x_m\;-\;f_jg_i\,(x_1\wedge\cdots\wedge x_m), \end{aligned}\tag{2.7}$$

em que $x^{\wedge}_{(k)}$ representa o vector $x_1\wedge\cdots\wedge x_{k-1}\wedge x_{k+1}\wedge\cdots\wedge x_m$ e a última igualdade em (2.7) é devida à ortonormalidade dos vectores $e_1,\ldots,e_n$. $\blacksquare$

O *operador número fermiónico no estado i* é o operador linear $N_i:\bigwedge^m\mathcal{H}\to\bigwedge^m\mathcal{H}$ definido por $N_i=f_ig_i$, $i=1,\ldots,n$. Dada a forma dos elementos da base (2.3) de $\bigwedge^m\mathcal{H}$, tem-se

$$N_i(e_{\alpha_1}\wedge\cdots\wedge e_{\alpha_m})\;=\;\begin{cases} e_{\alpha_1}\wedge\cdots\wedge e_{\alpha_m}, & \text{se}\quad i\in\{\alpha_1,\ldots,\alpha_m\}\\ 0, & \text{se}\quad i\notin\{\alpha_1,\ldots,\alpha_m\}\end{cases}.$$

Conclui-se que os valores próprios de $N_i:\bigwedge^m\mathcal{H}\to\bigwedge^m\mathcal{H}$ são o zero e o um, o que se traduz em termos físicos pelo princípio da exclusão de Pauli[1]. É evidente que N_i é um operador auto-adjunto e verifica-se $[N_i,N_j]=0$, $i,j=1,\ldots,n$.

[1] Não mais do que um fermião pode ocupar o mesmo estado quântico.

2.2.3 Transformação de Bogoliubov

Podem interpretar-se as álgebras simétrica $\Gamma^*(\mathcal{H})$ e de Grassmann $\Gamma^\wedge(\mathcal{H})$ como os espaços de estados de um sistema constituído por um número arbitrário, mas finito, de partículas, respectivamente, bosões e fermiões, se o espaço de estados de uma só partícula for o espaço de Hilbert $\mathcal{H}$.

Estendem-se, de modo óbvio, os operadores de criação e destruição bosónicos e fermiónicos à álgebra simétrica $\Gamma^*(\mathcal{H})$ e à álgebra de Grassmann $\Gamma^\wedge(\mathcal{H})$, respectivamente. Um estado de um sistema físico com m partículas é aplicado por um operador de criação num estado com $m+1$ partículas (cria-se uma partícula) e por um operador de destruição num estado com $m-1$ partículas (destroi-se uma partícula).

No decorrer deste capítulo, utilizaremos o símbolo $\epsilon = -1$ ($\epsilon = 1$) nas afirmações envolvendo operadores de criação e destruição de bosões (fermiões). A vantagem deste tratamento unificado é uma apresentação compacta de resultados paralelos.

Por conveniência, agrupam-se os operadores de destruição e criação definidos na álgebra simétrica (ou de Grassmann) sobre $\mathcal{H}$ num vector α de componentes

$$\alpha_i \;=\; g_i \qquad \text{e} \qquad \alpha_{n+i} \;=\; f_i, \qquad i = 1, \ldots, n. \tag{2.8}$$

Uma aplicação linear invertível que ao vector dos operadores de destruição e criação α associa um vector β de componentes

$$\beta_i \;=\; \tilde{g}_i \qquad \text{e} \qquad \beta_{n+i} \;=\; \tilde{f}_i, \qquad i = 1, \ldots, n. \tag{2.9}$$

e que preserva as relações de (anti-)comutação canónicas diz-se uma *transformação canónica*, vulgarmente conhecida por *transformação de Bogoliubov.* Apresenta-se, de seguida, uma caracterização útil de uma transformação de Bogoliubov.

Proposição 2.2.5 [33] *Sejam α e β os vectores coluna de componentes* (2.8) *e* (2.9)*, respectivamente. As afirmações seguintes são equivalentes:*

(i) *A aplicação que ao vector α associa o vector β é uma transformação de Bogoliubov;*

(ii) *A matriz T tal que $\beta = T\alpha$ satisfaz $TL_\epsilon T^{\mathrm{T}} = T^{\mathrm{T}} L_\epsilon T = L_\epsilon$, em que*

$$L_\epsilon \;=\; \begin{bmatrix} 0 & I_n \\ \epsilon I_n & 0 \end{bmatrix}. \tag{2.10}$$

Demonstração: Denotam-se por g e $\tilde{g}$ (ou f e $\tilde{f}$) os vectores coluna constituídos pelas primeiras (ou últimas) n componentes de α e β, respectivamente. Consideram-se as matrizes por blocos

$$M_\alpha = \begin{bmatrix} M_{gg} & M_{gf} \\ M_{fg} & M_{ff} \end{bmatrix} \qquad \text{e} \qquad M_\beta = \begin{bmatrix} M_{\tilde{g}\tilde{g}} & M_{\tilde{g}\tilde{f}} \\ M_{\tilde{f}\tilde{g}} & M_{\tilde{f}\tilde{f}} \end{bmatrix},$$

em que M_{gf} denota a matriz $n\times n$ de elemento genérico ij dado pelo comutador $[g_i, f_j]$, se $\epsilon=-1$, ou pelo anti-comutador $\{g_i, f_j\}$, se $\epsilon=1$, valendo analoga notação para M_{gg}, M_{fg}, M_{ff} e para as submatrizes de M_β. As relações de (anti-)comutação canónicas entre os operadores de destruição g_i e de criação f_i podem traduzir-se matricialmente pela igualdade $M_\alpha = L_\epsilon$. Seja

$$T = \begin{bmatrix} X & Y \\ U & V \end{bmatrix}, \qquad X, Y, U, V \in M_n,$$

tal que $\beta = T\alpha$, ou seja, $\tilde{g} = Xg+Yf$ e $\tilde{f} = Ug+Vf$. Mediante alguns cálculos, obtêm-se as seguintes relações:

$$\begin{aligned}
M_{\tilde{g}\tilde{g}} &= XM_{gf}Y^T + XM_{gg}X^T + YM_{ff}Y^T + YM_{fg}X^T, \\
M_{\tilde{g}\tilde{f}} &= XM_{gf}V^T + XM_{gg}U^T + YM_{ff}V^T + YM_{fg}U^T, \\
M_{\tilde{f}\tilde{f}} &= VM_{gf}U^T + VM_{gg}V^T + UM_{ff}U^T + UM_{fg}V^T, \\
M_{\tilde{f}\tilde{g}} &= VM_{gf}Y^T + VM_{gg}X^T + UM_{ff}Y^T + UM_{fg}X^T,
\end{aligned}$$

o que permite concluir que M_α e M_β se relacionam através da matriz T por

$$M_\beta = T\,M_\alpha\,T^T. \tag{2.11}$$

(i) $\Rightarrow$ (ii) : Se a transformação que aplica o vector α no vector β é de Bogoliubov, então preserva as relações de (anti-)comutação canónicas e $M_\beta = L_\epsilon$. De (2.11), resulta a condição $TL_\epsilon T^T = L_\epsilon$, donde se constata que $\epsilon L_\epsilon T^T L_\epsilon$ é a inversa à direita de T. Como esta coincide com a inversa à esquerda, então $\epsilon L_\epsilon T^T L_\epsilon T = I_{2n}$ e, multiplicando ambos os membros por L_ϵ, tem-se $T^T L_\epsilon T = L_\epsilon$.

(ii) $\Rightarrow$ (i) : Se a matriz T para a qual $\beta = T\alpha$ satisfaz $TL_\epsilon T^T = L_\epsilon$, por (2.11) e porque $M_\alpha = L_\epsilon$, tem-se $M_\beta = TM_\alpha T^T = L_\epsilon$, o que significa que os novos operadores $\tilde{g}_i$ e $\tilde{f}_i$ satisfazem as relações de (anti-)comutação canónicas, ou seja, a transformação que aplica o vector α no vector β é de Bogoliubov. Como, além disso, a matriz T satisfaz $T^T L_\epsilon T = L_\epsilon$, então T é invertível e pode escrever-se a relação inversa $\alpha = T^{-1}\beta$. ∎

Apesar da notação adoptada, os novos operadores $\tilde{g}_i$ não são necessariamente adjuntos dos operadores $\tilde{f}_i$. Tal ocorre quando a matriz T associada à transformação de Bogoliubov, considerada na Proposição 2.2.5 (ii), é uma matriz por blocos do tipo

$$T = \begin{bmatrix} X & Y \\ \overline{Y} & \overline{X} \end{bmatrix}, \qquad X, Y \in M_n. \tag{2.12}$$

Neste caso, a condição referida na Proposição 2.2.5 (ii) satisfeita por T pode substituir-se por $TJ_\epsilon T^* = T^* J_\epsilon T = J_\epsilon$, em que $J_\epsilon = I_n \oplus \epsilon I_n$, e

$$T^{-1} = \begin{bmatrix} X^* & \epsilon Y^{\mathrm{T}} \\ \epsilon Y^* & X^{\mathrm{T}} \end{bmatrix}.$$

Com vista a unificar resultados, fixamos a notação seguinte:

$$\mathbb{C}_{-1} = \{z \in \mathbb{C} : |z| < 1\} \qquad \text{e} \qquad \mathbb{C}_1 = \mathbb{C}.$$

Escrevemos simplesmente Γ^* e $\Gamma^\wedge$ para denotar, respectivamente, as álgebras simétrica e de Grassmann sobre $\mathbb{C}^2$.

Exemplo 2.2.1 Dado $z \in \mathbb{C}_\epsilon$, $\epsilon = -1$ ($\epsilon = 1$), sejam $\tilde{f}_i : \Gamma^* \to \Gamma^*$ ($\tilde{f}_i : \Gamma^\wedge \to \Gamma^\wedge$), $i = 1, 2$, os operadores lineares definidos por

$$\tilde{f}_1 = \frac{1}{\sqrt{1+\epsilon|z|^2}}\,(f_1 - \bar{z}g_2) \qquad \text{e} \qquad \tilde{f}_2 = \frac{1}{\sqrt{1+\epsilon|z|^2}}\,(f_2 + \epsilon\bar{z}g_1). \tag{2.13}$$

Os seus operadores adjuntos são dados por

$$\tilde{g}_1 = \frac{1}{\sqrt{1+\epsilon|z|^2}}\,(g_1 - zf_2) \qquad \text{e} \qquad \tilde{g}_2 = \frac{1}{\sqrt{1+\epsilon|z|^2}}\,(g_2 + \epsilon z f_1), \tag{2.14}$$

respectivamente. A transformação linear que aplica o vector (g_1, g_2, f_1, f_2) no vector $(\tilde{g}_1, \tilde{g}_2, \tilde{f}_1, \tilde{f}_2)$ é uma transformação de Bogoliubov associada a uma matriz por blocos do tipo (2.12), em que

$$X = \frac{1}{\sqrt{1+\epsilon|z|^2}}\, I_2 \qquad \text{e} \qquad Y = \frac{1}{\sqrt{1+\epsilon|z|^2}} \begin{bmatrix} 0 & -z \\ \epsilon z & 0 \end{bmatrix}. \tag{2.15}$$

Seja $\tilde{N}_i : \Gamma^*(\mathcal{H}) \to \Gamma^*(\mathcal{H})$ o operador linear definido por $\tilde{N}_i = \tilde{f}_i \tilde{g}_i$, $i = 1, \ldots, n$. A proposição seguinte é consequência directa das relações de comutação canónicas entre os operadores $\tilde{f}_i$ e $\tilde{g}_i$, $i = 1, \ldots, n$.

Proposição 2.2.6 *Se os operadores $\tilde{f}_i$ e $\tilde{g}_i$ satisfazem as relações de comutação canónicas, então também satisfazem as seguintes relações:*

$$[\tilde{N}_i, \tilde{f}_j^r] = r\,\delta_{ij}\,\tilde{f}_i^r \qquad e \qquad [\tilde{N}_i, \tilde{g}_j^r] = -r\,\delta_{ij}\,\tilde{g}_i^r, \qquad i, j = 1, \ldots, n, \quad r \in \mathbb{N}_0.$$

Demonstração: Seja $r \in \mathbb{N}_0$. Por indução sobre k, prova-se que

$$\tilde{N}_i\, \tilde{f}_j^r \;=\; k\, \delta_{ij}\, \tilde{f}_i^r + \tilde{f}_j^k\, \tilde{N}_i\, \tilde{f}_j^{r-k}, \qquad i,j = 1,\ldots,n, \qquad k = 0,\ldots,r. \tag{2.16}$$

De facto, (2.16) é trivial para $k=0$ e suponha-se (2.16) verdadeira para $k-1$. Então

$$\begin{aligned}
\tilde{N}_i\, \tilde{f}_j^r &= (k-1)\, \delta_{ij}\, \tilde{f}_i^r + \tilde{f}_j^{k-1}\, \tilde{N}_i\, \tilde{f}_j^{r-k+1} \\
&= (k-1)\, \delta_{ij}\, \tilde{f}_i^r + \tilde{f}_j^{k-1}\, \tilde{f}_i(\delta_{ij} + \tilde{f}_j\tilde{g}_i)\tilde{f}_j^{r-k} && (2.17) \\
&= (k-1)\, \delta_{ij}\, \tilde{f}_i^r + \tilde{f}_j^{k}(\delta_{ij} + \tilde{f}_i\tilde{g}_i)\tilde{f}_j^{r-k} && (2.18) \\
&= k\, \delta_{ij}\, \tilde{f}_i^r + \tilde{f}_j^k\, \tilde{N}_i\, \tilde{f}_j^{r-k},
\end{aligned}$$

em que (2.17) se deve às relações de comutação canónicas $[\tilde{g}_i, \tilde{f}_j] = \delta_{ij}$, e (2.18) resulta de $[\tilde{f}_i, \tilde{f}_j] = 0$ e de $\tilde{f}_i\, \delta_{ij} = \tilde{f}_j\, \delta_{ij}$. A igualdade (2.16) é, pois, válida e para $k=r$ origina o primeiro conjunto de relações pretendido. Por transconjugação destas relações, conclui-se a demonstração. ■

2.3 Operadores de Emparelhamento

Dada a álgebra simétrica sobre $\mathbb{C}^2$, Γ^*, o *operador de emparelhamento* $B : \Gamma^* \to \Gamma^*$ é o operador linear definido por

$$B \;=\; c\, f_1 g_1 + d\, f_2 g_2 + k\, f_1 f_2 + l\, g_1 g_2, \qquad c,d,k,l \in \mathbb{C}. \tag{2.19}$$

onde f_i e g_i, $i = 1,2$, são os operadores de criação e destruição de bosões. Dada a álgebra de Grassmann sobre $\mathbb{C}^2$, $\Gamma^\wedge$, define-se o *operador de emparelhamento* $B\colon \Gamma^\wedge \to \Gamma^\wedge$ como sendo o operador linear em (2.19), mas onde f_i e g_i, $i = 1,2$, são os operadores de criação e destruição de fermiões. Se necessário distinguir, o primeiro diz-se operador de emparelhamento bosónico e o segundo operador de emparelhamento fermiónico.

Seja $q \in \mathbb{Z}$. Para simplicidade de linguagem, utilizam-se as notações:

$$2\tau_q = |q| - q \qquad \text{e} \qquad 2\kappa_q = |q| + q.$$

Denota-se por $\Gamma^{(q)}$ o subespaço da álgebra simétrica sobre $\mathbb{C}^2$ gerado pelos vectores $e_1^{n+\tau_q} * e_2^{n+\kappa_q}$, $n \in \mathbb{N}_0$. Facilmente se verifica que estes subespaços $\Gamma^{(q)}$, $q \in \mathbb{Z}$, permitem decompor Γ^* numa soma directa.

Proposição 2.3.1 *A álgebra simétrica sobre $\mathbb{C}^2$ admite a seguinte decomposição:*

$$\Gamma^* \;=\; \bigoplus_{q=-\infty}^{+\infty} \Gamma^{(q)}.$$

Demonstração: Os conjuntos $\Gamma^{(q)}$, $q \in \mathbb{Z}$, são dois a dois disjuntos pelo que a sua reunião é uma soma directa. O m-ésimo espaço simétrico sobre $\mathbb{C}^2$ tem dimensão $m+1$. Se o número de ocupação de nível 1 na sucessão $\alpha \in G_{m,2}$ é o número inteiro não negativo k não superior a m, então o número de ocupação de nível 2 é $m-k$ e $\{e_1^k * e_2^{m-k} : \ k = 0, \ldots, m\}$ é uma base de $\mathbb{C}^2_{(m)}$. Assim, o conjunto de todos os vectores $e_1^k * e_2^{m-k}$, $k = 0, \ldots, m$, $m \in \mathbb{N}_0$, é um conjunto gerador da álgebra Γ^*. Se $m \geq 2k$ ($m < 2k$), estes geradores são elementos de um subespaço $\Gamma^{(q)}$, para $q \geq 0$ ($q < 0$). Confirma-se, assim, a inclusão de Γ^* na reunião de todos os subespaços $\Gamma^{(q)}$, $q \in \mathbb{Z}$.

No que se refere à verificação da inclusão contrária, basta mostrar que os geradores de todos os subespaços $\Gamma^{(q)}$ pertencem a $\mathbb{C}^2_{(m)}$, para um determinado m, o que efectivamente se verifica, pois o vector $e_1^{n+\tau_q} * e_2^{n+\kappa_q}$, gerador de $\Gamma^{(q)}$, é um elemento de $\mathbb{C}^2_{(2n+|q|)}$, para todo $n \in \mathbb{N}_0$ e $q \in \mathbb{Z}$. ■

Os subespaços $\Gamma^{(q)}$, $q \in \mathbb{Z}$, que decompõem a álgebra simétrica sobre $\mathbb{C}^2$ numa soma directa, satisfazem a seguinte propriedade, que permite justificar o termo emparelhamento adoptado para designar o operador $B : \Gamma^* \to \Gamma^*$ em (2.19).

Proposição 2.3.2 *Dado $q \in \mathbb{Z}$, o subespaço $\Gamma^{(q)}$ é invariante pelo operador de emparelhamento bosónico $B : \Gamma^* \to \Gamma^*$ definido por* (2.19).

Demonstração: Fixando $q \in \mathbb{Z}$, verifica-se que as imagens por $B : \Gamma^* \to \Gamma^*$ dos vectores geradores de $\Gamma^{(q)}$ pertencem a $\Gamma^{(q)}$. De facto,

$$\begin{aligned} B(e_1^{n+\tau_q} * e_2^{n+\kappa_q}) &= (n(c+d) + \tau_q c + \kappa_q d)\, e_1^{n+\tau_q} * e_2^{n+\kappa_q} + k\, e_1^{n+\tau_q+1} * e_2^{n+\kappa_q+1} \\ &+ ln\,(n+|q|)\, e_1^{n+\tau_q-1} * e_2^{n+\kappa_q-1} \in \Gamma^{(q)}, \qquad n \in \mathbb{N}_0. \end{aligned}$$

Sendo B um operador linear, satisfaz $B(\Gamma^{(q)}) \subseteq \Gamma^{(q)}$, para cada $q \in \mathbb{Z}$. ■

Podemos munir cada subespaço $\Gamma^{(q)}$, $q \in \mathbb{Z}$, de um produto interno. Considerando

$$\varphi_i = \sum_{n=0}^{+\infty} c_{in}\, e_1^{n+\tau_q} * e_2^{n+\kappa_q}, \qquad c_{in} \in \mathbb{C}, \quad n \in \mathbb{N}_0, \qquad i = 1, 2,$$

dois elementos arbitrários de $\Gamma^{(q)}$, o seu produto interno, induzido pelo produto interno

$$\langle e_1^{n+\tau_q} * e_2^{n+\kappa_q}, e_1^{n+\tau_q} * e_2^{n+\kappa_q} \rangle_{2n+|q|} = (n+\tau_q)!\,(n+\kappa_q)!,$$

é dado por

$$\langle \varphi_1, \varphi_2 \rangle = \sum_{n=0}^{+\infty} c_{1n}\, \overline{c}_{2n}\, (n+\tau_q)!\,(n+\kappa_q)!.$$

Atendendo à Proposição 2.3.2, pode considerar-se restrito ao subespaço $\Gamma^{(q)}$, $q \in \mathbb{Z}$, o operador de emparelhamento bosónico B definido por (2.19). A sua representação matricial, na base canónica de $\Gamma^{(q)}$,

$$\mathcal{B}^q = \left\{ \frac{e_1^{n+\tau_q} * e_2^{n+\kappa_q}}{\sqrt{(n+\tau_q)!\,(n+\kappa_q)!}} : n \in \mathbb{N}_0 \right\},$$

é a matriz tridiagonal infinita

$$T_B^q = \begin{bmatrix} a_0 & b_1 & 0 & 0 & \dots \\ c_1 & a_1 & b_2 & 0 & \dots \\ 0 & c_2 & a_2 & b_3 & \dots \\ 0 & 0 & c_3 & a_3 & \ddots \\ \vdots & \vdots & \vdots & \ddots & \ddots \end{bmatrix}, \tag{2.20}$$

em que

$$\begin{aligned} a_n &= n(c+d) + \tau_q c + \kappa_q d, && n \in \mathbb{N}_0, \\ b_n &= l\sqrt{n(n+|q|)}, && n \in \mathbb{N}, \\ c_n &= k\sqrt{n(n+|q|)}, && n \in \mathbb{N}. \end{aligned} \tag{2.21}$$

2.3.1 Propriedades Espectrais

O estudo das propriedades espectrais de alguns operadores de emparelhamento é o nosso próximo objectivo. Para simplificar esse estudo, utiliza-se uma transformação de Bogoliubov adequada, mais precisamente a introduzida no Exemplo 2.2.1. Denota-se por ι a aplicação identidade em Γ^* ou $\Gamma^\wedge$. Escreve-se ι^* para especificar que a aplicação identidade é definida em Γ^*, ou num seu subespaço $\Gamma^{(q)}$, $q \in \mathbb{Z}$.

Proposição 2.3.3 *A transformação de Bogoliubov determinada por* (2.13) *e* (2.14)*, com $z \in \mathbb{C}_\epsilon$, $\epsilon = -1$ ($\epsilon = 1$), aplica o operador de emparelhamento $B : \Gamma^* \to \Gamma^*$ ($B : \Gamma^\wedge \to \Gamma^\wedge$) definido por* (2.19) *na forma*

$$B = \lambda_0\, \iota + \tilde{c}\, \tilde{f}_1 \tilde{g}_1 + \tilde{d}\, \tilde{f}_2 \tilde{g}_2 + \tilde{k}\, \tilde{f}_1 \tilde{f}_2 + \tilde{l}\, \tilde{g}_1 \tilde{g}_2, \tag{2.22}$$

onde

$$\lambda_0 = \frac{1}{1+\epsilon|z|^2}\left((c+d)|z|^2 + k\bar{z} - \epsilon l z\right), \tag{2.23}$$

$$\tilde{c} = \frac{1}{1+\epsilon|z|^2}\left(c - \epsilon d|z|^2 - \epsilon k\bar{z} + l z\right), \tag{2.24}$$

$$\tilde{d} = \frac{1}{1+\epsilon|z|^2}\left(-\epsilon c|z|^2 + d - \epsilon k\bar{z} + l z\right), \tag{2.25}$$

$$\tilde{k} = \frac{1}{1+\epsilon|z|^2}\left((c+d)z + k + lz^2\right), \tag{2.26}$$

$$\tilde{l} = \frac{1}{1+\epsilon|z|^2}\left(-\epsilon(c+d)\bar{z} + k\bar{z}^2 + l\right). \tag{2.27}$$

Além disso,

$$\tilde{c} = c - \epsilon\lambda_0 \qquad e \qquad \tilde{d} = d - \epsilon\lambda_0. \tag{2.28}$$

Demonstração: A transformação de Bogoliubov que aplica o vector $\alpha^T = (g_1, g_2, f_1, f_2)$ no vector $\beta^T = (\tilde{g}_1, \tilde{g}_2, \tilde{f}_1, \tilde{f}_2)$, como no Exemplo 2.2.1, está associada a uma matriz por blocos T do tipo (2.12), de submatrizes X e Y dadas por (2.15). Como $\alpha = T^{-1}\beta$ e

$$T^{-1} = \frac{1}{\sqrt{1+\epsilon|z|^2}}\begin{bmatrix} 1 & 0 & 0 & z \\ 0 & 1 & -\epsilon z & 0 \\ 0 & \bar{z} & 1 & 0 \\ -\epsilon\bar{z} & 0 & 0 & 1 \end{bmatrix},$$

obtêm-se as seguintes relações inversas:

$$f_1 = \frac{1}{\sqrt{1+\epsilon|z|^2}}(\tilde{f}_1 + \bar{z}\tilde{g}_2), \qquad f_2 = \frac{1}{\sqrt{1+\epsilon|z|^2}}(\tilde{f}_2 - \epsilon\bar{z}\tilde{g}_1) \tag{2.29}$$

e

$$g_1 = \frac{1}{\sqrt{1+\epsilon|z|^2}}(\tilde{g}_1 + z\tilde{f}_2), \qquad g_2 = \frac{1}{\sqrt{1+\epsilon|z|^2}}(\tilde{g}_2 - \epsilon z\tilde{f}_1). \tag{2.30}$$

Substituindo f_i e g_i, $i = 1, 2$, no operador $B = cf_1g_1 + df_2g_2 + kf_1f_2 + lg_1g_2$, pelas expressões (2.29) e (2.30), tem-se o pretendido. ∎

Centramos primeiro a atenção em operadores de emparelhamento auto-adjuntos.

Proposição 2.3.4 *O operador de emparelhamento B dado por* (2.19) *é auto-adjunto se e só se $c, d \in \mathbb{R}$ e $l = -\epsilon\bar{k}$.*

Demonstração: Sejam $c, d, k, l \in \mathbb{C}$. Atendendo às relações de (anti-)comutação canónicas $f_2f_1 + \epsilon f_1f_2 = g_2g_1 + \epsilon g_1g_2 = 0$, o operador adjunto do operador de emparelhamento $B = cf_1g_1 + df_2g_2 + kf_1f_2 + lg_1g_2$ é dado por $B^* = \bar{c}f_1g_1 + \bar{d}f_2g_2 - \epsilon\bar{l}f_1f_2 - \epsilon\bar{k}g_1g_2$. Assim, B é auto-adjunto se e só se $c = \bar{c}, d = \bar{d}$ e $l = -\epsilon\bar{k}$. ∎

Conclui-se, de imediato, da Proposição 2.3.4 que o operador de emparelhamento B na forma (2.22) é auto-adjunto se e só se $\tilde{c}, \tilde{d} \in \mathbb{R}$ e $\tilde{l} = -\epsilon\bar{\tilde{k}}$.

De agora em diante, para simplicidade de linguagem, consideramos

$$\Delta_B = (c+d)^2 - 4kl \qquad \text{e} \qquad \zeta_{c+d} = \begin{cases} e^{i\arg(c+d)}, & \text{se } c+d \neq 0 \\ 0, & \text{se } c+d = 0 \end{cases}.$$

Além disso, sejam $\tilde{N}_i = \tilde{f}_i \tilde{g}_i$, $i = 1, 2$, em que $\tilde{f}_i$ e $\tilde{g}_i$ são os operadores lineares definidos na álgebra simétrica Γ^* ou na álgebra de Grassmann $\Gamma^\wedge$ sobre $\mathbb{C}^2$, como no Exemplo 2.2.1.

Proposição 2.3.5 *Se o operador de emparelhamento* $B : \Gamma^\wedge \to \Gamma^\wedge$ $(B : \Gamma^* \to \Gamma^*)$ *definido em* (2.19) *é auto-adjunto* (e $\Delta_B > 0$ *ou* $k = 0$)*, então pode reduzir-se, por meio de uma transformação de Bogoliubov, à forma*

$$B = \lambda_0\, \iota + (c - \epsilon\lambda_0)\, \tilde{N}_1 + (d - \epsilon\lambda_0)\, \tilde{N}_2,$$

com $\epsilon = 1$ *e* $\lambda_0 = \frac{1}{2}(c+d) \pm \frac{1}{2}\sqrt{\Delta_B}$ $\left(\epsilon = -1 \text{ e } \lambda_0 = -\frac{1}{2}(c+d) + \frac{1}{2}\,\zeta_{c+d}\sqrt{\Delta_B}\right)$.

Demonstração: Pela Proposição 2.3.4, verifica-se que $c, d \in \mathbb{R}$ e $l = -\epsilon\bar{k}$. Pela Proposição 2.3.3, por meio de uma transformação de Bogoliubov adequada, pode escrever-se B na forma (2.22), com λ_0, $\tilde{c}$, $\tilde{d}$ e $\tilde{k}$ dados por (2.23), (2.24), (2.25) e (2.26), respectivamente, e $\tilde{l} = -\epsilon\bar{\tilde{k}}$.

Seja $\Delta_B = 0$. Se $B : \Gamma^\wedge \to \Gamma^\wedge$, então $c + d = k = 0$. Se $B : \Gamma^* \to \Gamma^*$ e $k = 0$, então $c + d = 0$. Em ambos os casos, o resultado é trivial, já que $\tilde{c} = c$, $\tilde{d} = d$ e $\lambda_0 = \tilde{k} = 0$, qualquer que seja o $z \in \mathbb{C}_\epsilon$.

Seja, agora, $\Delta_B > 0$. Se $B : \Gamma^\wedge \to \Gamma^\wedge$ ou $B : \Gamma^* \to \Gamma^*$, é possível determinar $z \in \mathbb{C}_\epsilon$, tal que $\tilde{k} = 0$. De facto, podemos escolher uma solução z da equação quadrática

$$-\epsilon\bar{k}z^2 + (c+d)\, z + k \;=\; 0, \tag{2.31}$$

para a qual $\tilde{k}$ se anula. A escolha é a seguinte. Para $k = 0$, como $c + d \neq 0$, toma-se $z = 0$. Para $k \neq 0$, tem-se

$$z \;=\; \frac{\epsilon(c+d) \pm \sqrt{\Delta_B}}{2\bar{k}}, \tag{2.32}$$

e o produto destas duas raízes, $-\epsilon k/\bar{k}$, tem módulo um. Portanto, se $\epsilon = -1$, uma das raízes z pertence a $\mathbb{C}_{-1}$. Se $\epsilon = 1$, então $\tilde{k} = 0$, para qualquer uma destas raízes. Assim, basta centrar a atenção no operador $B = \lambda_0\, \iota + \tilde{c}\, \tilde{N}_1 + \tilde{d}\, \tilde{N}_2$. De (2.24) e (2.25), vem

$$\tilde{c} + \tilde{d} \;=\; \frac{(c+d)\,(1 - \epsilon|z|^2) - 2\epsilon k\bar{z} - 2\epsilon\bar{k}z}{1 + \epsilon|z|^2}.$$

Atendendo a (2.32), mediante alguns cálculos, tem-se $\tilde{c} + \tilde{d} = \mp\,\epsilon\sqrt{\Delta_B}$. Por outro lado, de (2.28), vem $\tilde{c} + \tilde{d} = c + d - \epsilon 2\lambda_0$. Logo, $2\lambda_0 = \epsilon(c+d) \pm \sqrt{\Delta_B}$, obtendo-se o resultado para $\epsilon = 1$. Se $\epsilon = -1$ e $c+d > 0$, considera-se o sinal $'+'$ em vez do sinal $'\pm'$ na expressão (2.32), de modo a que z pertença a $\mathbb{C}_{-1}$. Se $\epsilon = -1$ e $c + d < 0$, substitui-se o sinal $'\pm'$ na expressão (2.32) pelo sinal $'-'$, caso contrário $z \notin \mathbb{C}_{-1}$. ■

Nota 2.3.1 Perante o operador de emparelhamento auto-adjunto bosónico B, a restrição $\Delta_B > 0$ ou $k = 0$ imposta na Proposição 2.3.5, é essencial. Se $\Delta_B \leq 0$ e $k \neq 0$, verifica-se facilmente que ambas as raízes de (2.31) têm módulo um e, por isso, não se pode escolher $z \in \mathbb{C}_{-1}$, tal que $\tilde{k} = 0$.

Proposição 2.3.6 *Se o operador de emparelhamento $B : \Gamma^* \to \Gamma^*$ definido em* (2.19) *é auto-adjunto, um número complexo z satisfaz*

$$[B, g_1 - zf_2] \;=\; \frac{1}{2}\left(d - c \pm \sqrt{\Delta_B}\right)(g_1 - zf_2) \tag{2.33}$$

ou

$$[B, g_2 - zf_1] \;=\; \frac{1}{2}\left(c - d \pm \sqrt{\Delta_B}\right)(g_2 - zf_1) \tag{2.34}$$

se e só se z é uma raiz de (2.31)*, com $\epsilon = -1$.*

Demonstração: A implicação ($\Leftarrow$) é imediata. Suponhamos que $z \in \mathbb{C}$ satisfaz (2.33). Utilizando as relações de comutação canónicas e o facto de $l = \bar{k}$, pois B é auto-adjunto, facilmente se verifica que

$$[B, g_1 - zf_2] \;=\; -\left(c + \bar{k}z\right) g_1 - (k + dz)\, f_2. \tag{2.35}$$

Mostraremos a existência de $w \in \mathbb{C}$ tal que $[B, g_1 - zf_2] \;=\; w(g_1 - zf_2)$, o que equivale, por (2.35), a determinar $w \in \mathbb{C}$ tal que

$$\begin{bmatrix} c & \bar{k} \\ k & d \end{bmatrix} \begin{bmatrix} 1 \\ z \end{bmatrix} \;=\; w \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ z \end{bmatrix}. \tag{2.36}$$

As soluções w de (2.36) satisfazem

$$\det \begin{bmatrix} w + c & \bar{k} \\ -k & w - d \end{bmatrix} \;=\; 0,$$

ou seja, são tais que $2w = d - c \pm \sqrt{\Delta_B}$. A partir de (2.36), tem-se $z = -(c + w)/\bar{k}$, isto é, z é uma raiz da equação quadrática (2.31), com $\epsilon = -1$. Se $z \in \mathbb{C}$ satisfaz (2.34), a prova decorre de modo análogo. ■

Proposição 2.3.7 *Sejam $f_i, g_i : \Gamma^* \to \Gamma^*$, $i = 1, 2$, e $z \in \mathbb{C}$. Existe um vector não nulo $u \in \Gamma^*$ satisfazendo $(g_1 - zf_2)\, u = 0$ e $(g_2 - zf_1)\, u = 0$ se e só se $|z| < 1$. Esse vector u é dado pela fórmula*

$$u \;=\; \sum_{n=0}^{+\infty} c_0 \,\frac{z^n}{n!}\, f_1^n f_2^n(1), \qquad c_0 \in \mathbb{C} \setminus \{0\}. \tag{2.37}$$

Demonstração: Considerando um elemento arbitrário $u = \sum_{n,m=0}^{+\infty} c_{nm} f_1^n f_2^m(1) \in \Gamma^*$, $c_{nm} \in \mathbb{C} \setminus \{0\}$, facilmente se verifica que

$$\begin{aligned}(g_1 - zf_2)\, u &= \sum_{n,m=0}^{+\infty} \left(c_{n+1m+1}(n+1) - c_{nm}z\right)\, f_1^n f_2^{m+1}(1), \qquad (2.38)\\ (g_2 - zf_1)\, u &= \sum_{n,m=0}^{+\infty} \left(c_{n+1m+1}(m+1) - c_{nm}z\right)\, f_1^{n+1} f_2^{m}(1). \qquad (2.39)\end{aligned}$$

Por hipótese, são nulos os vectores apresentados em (2.38) e (2.39), pelo que

$$c_{n+1m+1}(n+1) \;=\; c_{nm}z \;=\; c_{n+1m+1}(m+1), \qquad n, m \in \mathbb{N}_0, \qquad (2.40)$$

donde se conclui que $(n-m)\, c_{n+1m+1} = 0$, ou seja, $c_{nm} = c_n \delta_{nm}$, $n, m \in \mathbb{N}_0$. Portanto,

$$u \;=\; \sum_{n=0}^{+\infty} c_n\, f_1^n f_2^n(1) \;\in\; \Gamma^{(0)}.$$

De novo, a partir de (2.40), com $m = n$, vem $c_{n+1}(n+1) = c_n z$, $n \in \mathbb{N}_0$. Por indução sobre n, mostra--se que $c_n = c_0\, z^n/n!$, $n \in \mathbb{N}_0$, com $c_0 \in \mathbb{C} \setminus \{0\}$, o que confirma ser (2.37) a expressão do vector u. Ora, o vector u pertence ao espaço Γ^* se e só se $|z| < 1$. ■

Doravante, dado $z \in \mathbb{C}_{-1}$, escreveremos simplesmente $\mathrm{e}^{zf_1f_2}(1)$ para abreviar a fórmula $\sum_{n=0}^{+\infty} \frac{z^n}{n!}\, f_1^n f_2^n(1)$, interveniente na Proposição 2.3.7.

Proposição 2.3.8 *Se o operador de emparelhamento $B : \Gamma^* \to \Gamma^*$ definido em* (2.19) *é auto-adjunto e se B admite um vector próprio $u \in \Gamma^*$ associado a um valor próprio real λ, então $\Delta_B \geq 0$.*

Demonstração: Suponhamos que $\Delta_B < 0$. Então $k \neq 0$. Pela Proposição 2.3.6, um número complexo z satisfaz (2.33) ou (2.34) se e só se z é uma raiz de (2.31), com $\epsilon = -1$. Como observado na Nota 2.3.1, a hipótese $\Delta_B < 0$ implica que $|z| = 1$. Mediante alguns cálculos,

$$B(g_1 - zf_2)u = (g_1 - zf_2)Bu + [B, g_1 - zf_2]u = \Big(\lambda + \frac{1}{2}(d-c) + \frac{i}{2}\sqrt{-\Delta_B}\Big)(g_1 - zf_2)u.$$

Então, $(g_1 - zf_2)u$ anula-se ou é um vector próprio de B associado ao valor próprio $\lambda + \frac{1}{2}(d-c) + \frac{i}{2}\sqrt{-\Delta_B}$. Como um operador auto-adjunto não tem valores próprios complexos, esta última hipótese não se verifica e $(g_1 - zf_2)u = 0$. De modo similar, a partir de

$$B(g_2 - zf_1)u = \Big(\lambda + \frac{1}{2}(c-d) + \frac{i}{2}\sqrt{-\Delta_B}\Big)\,(g_2 - zf_1)u,$$

conclui-se que $(g_2 - zf_1)u = 0$. Pela Proposição 2.3.7, as condições $(g_1 - zf_2)\,u = 0$ e $(g_2 - zf_1)\,u = 0$ implicam que $|z| < 1$, uma contradição. Logo, $\Delta_B \geq 0$. ■

Restringimos, agora, o estudo aos operadores $\tilde{N}_1$ e $\tilde{N}_2$ definidos na álgebra simétrica sobre $\mathbb{C}^2$ à custa dos operadores $\tilde{f}_i$ e $\tilde{g}_i$, $i = 1, 2$, do Exemplo 2.2.1.

Proposição 2.3.9 *Os valores próprios dos operadores $\tilde{N}_i : \Gamma^* \to \Gamma^*$, $i = 1, 2$, são os números inteiros não negativos e os vectores próprios comuns correspondentes aos valores próprios n_1 e n_2 de $\tilde{N}_1$ e $\tilde{N}_2$, respectivamente, são da forma*

$$v_{n_1 n_2} \;=\; c_0 \tilde{f}_1^{\,n_1} \tilde{f}_2^{\,n_2}\, \mathrm{e}^{z f_1 f_2}(1), \qquad c_0 \in \mathbb{C} \setminus \{0\}, \qquad z \in \mathbb{C}_{-1}. \tag{2.41}$$

Demonstração: Como os operadores $\tilde{N}_1$ e $\tilde{N}_2$ comutam entre si, então têm vectores próprios em comum. Seja $u \in \Gamma^*$ um vector não nulo, tal que $\tilde{N}_i u = \lambda_i u$, $i = 1, 2$. Da Proposição 2.2.6, com $r = 1$, vem $\tilde{N}_i \tilde{g}_i u = (\lambda_i - 1)\, \tilde{g}_i u$, $i = 1, 2$, o que significa que ou $\tilde{g}_i u = 0$ ou $\tilde{g}_i u$ é um vector próprio de $\tilde{N}_i$ associado ao valor próprio $\lambda_i - 1$, $i = 1, 2$.

Se $\tilde{g}_1 u = \tilde{g}_2 u = 0$, então os valores próprios de $\tilde{N}_1$ e $\tilde{N}_2$ são ambos nulos. Pela Proposição 2.3.7, conclui-se que u é da forma $c_0\, \mathrm{e}^{z f_1 f_2}(1)$, com $c_0 \in \mathbb{C} \setminus \{0\}$ e $z \in \mathbb{C}_{-1}$, isto é, vale (2.41) com $n_1 = n_2 = 0$.

Se $\tilde{g}_1 u \neq 0$ ou $\tilde{g}_2 u \neq 0$, repete-se o procedimento antes descrito. De facto, existem $n_1, n_2 \in \mathbb{Z}$, tais que $\tilde{g}_1^{\,n_1+1} \tilde{g}_2^{\,n_2} u = \tilde{g}_1^{\,n_1} \tilde{g}_2^{\,n_2+1} u = 0$ e $v = \tilde{g}_1^{\,n_1} \tilde{g}_2^{\,n_2} u \neq 0$ é um vector próprio comum a $\tilde{N}_1$ e $\tilde{N}_2$. Sendo $\tilde{N}_1$ e $\tilde{N}_2$ operadores semi-definidos positivos, os seus valores próprios associados ao vector próprio v, $\lambda_1 - n_1$ e $\lambda_2 - n_2$, respectivamente, são não negativos. Este procedimento pára, quando $\lambda_1 = n_1$ e $\lambda_2 = n_2$. Atendendo à expressão (2.14) dos operadores $\tilde{g}_1$ e $\tilde{g}_2$, no Exemplo 2.2.1, e como $\tilde{g}_1 v = \tilde{g}_2 v = 0$, pela Proposição 2.3.7, tem-se $v = d_0\, \mathrm{e}^{z f_1 f_2}(1)$, com $d_0 \in \mathbb{C} \setminus \{0\}$ e $z \in \mathbb{C}_{-1}$. Ora, $v = \tilde{g}_1^{\,n_1} \tilde{g}_2^{\,n_2} u$ implica $\tilde{f}_2^{\,n_2} v = n_2! \tilde{g}_1^{\,n_1} u$, logo $\tilde{f}_1^{\,n_1} \tilde{f}_2^{\,n_2} v = n_1! n_2!\, u$, donde se conclui o pretendido. ■

Investigam-se, de seguida, os valores próprios e vectores próprios de operadores de emparelhamento auto-adjuntos, restritos ao subespaço $\Gamma^{(0)}$.

Teorema 2.3.1 *Se o operador de emparelhamento bosónico B, restrito a $\Gamma^{(0)}$, definido por* (2.19) *é auto-adjunto e $\Delta_B > 0$, então os seus valores próprios são dados por*

$$\lambda_n \;=\; -\frac{1}{2}(c + d) + \frac{2n + 1}{2}\, \zeta_{c+d} \sqrt{\Delta_B}, \qquad n \in \mathbb{N}_0. \tag{2.42}$$

Os vectores próprios de B associados ao valor próprio λ_n são os vectores

$$v_n \;=\; c_0 \tilde{f}_1^{\,n} \tilde{f}_2^{\,n}\, \mathrm{e}^{z f_1 f_2}(1), \qquad c_0 \in \mathbb{C} \setminus \{0\},$$

em que z é a raiz da equação quadrática (2.31) *em $\mathbb{C}_{-1}$.*

Demonstração: Pela Proposição 2.3.5, por meio da transformação de Bogoliubov determinada por (2.13) e (2.14), com $z \in \mathbb{C}_{-1}$ raiz da equação quadrática (2.31), pode reduzir-se o operador de emparelhamento auto-adjunto B à forma

$$B \;=\; \lambda_0\, \iota^* + (c+\lambda_0)\tilde{N}_1 + (d+\lambda_0)\tilde{N}_2,$$

em que $\lambda_0 = -\frac{1}{2}(c+d) + \frac{1}{2}\,\zeta_{c+d}\sqrt{\Delta_B}$. Facilmente se verifica que os operadores $\tilde{N}_1 - \tilde{N}_2$ e $N_1 - N_2$ coincidem em Γ^* e, assim, os operadores $\tilde{N}_1$ e $\tilde{N}_2$ são iguais quando restritos a $\Gamma^{(0)}$. Pela Proposição 2.3.9, os seus valores próprios são os números inteiros não negativos. Sendo $B - \lambda_0\, \iota^*$ uma combinação linear dos operadores $\tilde{N}_1$ e $\tilde{N}_2$ que comutam entre si, resulta da Proposição 2.3.9 que os valores próprios do operador B restrito a $\Gamma^{(0)}$ são $\lambda_n = \lambda_0 + (c+d+2\lambda_0)\, n$, $n \in \mathbb{N}_0$, e vale (2.42). Os vectores próprios comuns a $\tilde{N}_1$ e a $\tilde{N}_2$ são os vectores próprios de B e, pela Proposição 2.3.9, tem-se o resultado. ∎

De modo análogo, provam-se as seguintes propriedades espectrais relativas a operadores de emparelhamento auto-adjuntos, restritos ao subespaço $\Gamma^{(q)}$, para qualquer $q \in \mathbb{Z}$.

Teorema 2.3.2 *Se o operador de emparelhamento bosónico B, restrito a $\Gamma^{(q)}$, $q \in \mathbb{Z}$, definido por* (2.19)*, é auto-adjunto e $\Delta_B > 0$, então os seus valores próprios são*

$$\lambda_{nq} \;=\; \frac{1+|q|}{2}\left(-(c+d) + \zeta_{c+d}\sqrt{\Delta_B}\right) + n\,\zeta_{c+d}\sqrt{\Delta_B} + \tau_q c + \kappa_q d, \qquad n \in \mathbb{N}_0.$$

Os vectores próprios de B associados ao valor próprio λ_{nq} são da forma

$$v_{nq} = c_0 \tilde{f}_1^{\,n+\tau_q} \tilde{f}_2^{\,n+\kappa_q}\, \mathrm{e}^{z f_1 f_2}(1), \qquad c_0 \in \mathbb{C}\setminus\{0\},$$

em que z é a raiz da equação quadrática (2.31) *em $\mathbb{C}_{-1}$.*

Mais geralmente, caracterizam-se os valores próprios e vectores próprios dos operadores de emparelhamento auto-adjuntos definidos na álgebra simétrica sobre $\mathbb{C}^2$.

Teorema 2.3.3 *Se o operador de emparelhamento bosónico $B : \Gamma^* \to \Gamma^*$, definido por* (2.19)*, é auto-adjunto e $\Delta_B > 0$, então os seus valores próprios são*

$$\lambda_{n_1 n_2} \;=\; \frac{1}{2}\left(-(c+d) + \zeta_{c+d}\sqrt{\Delta_B}\right)(n_1+n_2+1) + c n_1 + d n_2, \qquad n_1, n_2 \in \mathbb{N}_0. \quad (2.43)$$

Os vectores próprios de B associados ao valor próprio $\lambda_{n_1 n_2}$ são

$$v_{n_1 n_2} \;=\; c_0\, \tilde{f}_1^{\,n_1} \tilde{f}_2^{\,n_2}\, \mathrm{e}^{z f_1 f_2}(1), \qquad c_0 \in \mathbb{C}\setminus\{0\},$$

em que z é a raiz da equação quadrática (2.31) *em $\mathbb{C}_{-1}$.*

Demonstração: Pela Proposição 2.3.5, pode reduzir-se o operador de emparelhamento auto-adjunto B, por meio de uma transformação de Bogoliubov adequada, à forma

$$B \;=\; \lambda_0\, \iota^* + (c+\lambda_0)\tilde{N}_1 + (d+\lambda_0)\tilde{N}_2,$$

em que $\lambda_0 = -\frac{1}{2}(c+d) + \frac{1}{2}\,\zeta_{c+d}\sqrt{\Delta_B}$. Conhecidos os valores próprios dos operadores $\tilde{N}_i$, $i = 1, 2$, caracterizados na Proposição 2.3.9, obtêm-se os valores próprios do operador B, $\lambda_{n_1 n_2} = \lambda_0(n_1 + n_2 + 1) + cn_1 + dn_2$, $n_1, n_2 \in \mathbb{N}_0$, isto é, vale (2.43). Os vectores próprios comuns a $\tilde{N}_1$ e $\tilde{N}_2$, associados aos valores próprios n_1 e n_2, respectivamente, são os vectores próprios do operador B. Pela Proposição 2.3.9, conclui-se o pretendido. ∎

2.3.2 Contradomínio Numérico no Caso Bosónico

A caracterização do contradomínio numérico do operador de emparelhamento bosónico B definido em (2.19), restrito ao subespaço $\Gamma^{(q)}$, $q \in \mathbb{Z}$, é o principal objectivo desta subsecção. Apresenta-se no Lema 2.3.1 uma relação de inclusão para $W(B\,|_{\Gamma^{(q)}})$. Este lema será usado nas demonstrações dos teoremas principais desta subsecção.

Lema 2.3.1 *Sejam $B : \Gamma^* \to \Gamma^*$ o operador de emparelhamento definido em* (2.19) *e*

$$W \;=\; \left\{ \frac{(c+d)|z|^2 + k\overline{z} + lz}{1-|z|^2} \;:\; z \in \mathbb{C}_{-1} \right\}. \tag{2.44}$$

Então $(1+|q|)\,W + \tau_q c + \kappa_q d \subseteq W(B\,|_{\Gamma^{(q)}})$, qualquer que seja $q \in \mathbb{Z}$.

Demonstração: Seja $q \in \mathbb{Z}$. Considerando um elemento arbitrário $\psi \in \Gamma^{(q)}$,

$$\psi \;=\; \sum_{n=0}^{+\infty} c_n\, e_1^{n+\tau_q} * e_2^{n+\kappa_q}, \qquad c_n \in \mathbb{C},$$

verifica-se o seguinte, para $q \geq 0$,

$$\begin{aligned}
\langle \psi, \psi \rangle &= \sum_{n=0}^{+\infty} |c_n|^2 n!(n+q)!, \\
\langle f_1 f_2 \psi, \psi \rangle &= \sum_{n=0}^{+\infty} c_n\, \bar{c}_{n+1}\, (n+1)!(n+q+1)!, \\
\langle g_1 g_2 \psi, \psi \rangle &= \sum_{n=0}^{+\infty} c_{n+1} \bar{c}_n\, (n+1)!(n+q+1)!, \\
\langle f_1 g_1 \psi, \psi \rangle &= \sum_{n=0}^{+\infty} n\, |c_n|^2\, n!(n+q)!, \\
\langle f_2 g_2 \psi, \psi \rangle &= \sum_{n=0}^{+\infty} (n+q)\, |c_n|^2\, n!(n+q)!.
\end{aligned}$$

Se $c_n = z^n/n!$, $z \in \mathbb{C}_{-1}$, as séries antes consideradas convergem e tem-se

$$\begin{aligned}
\langle \psi, \psi \rangle &= \sum_{n=0}^{+\infty} \prod_{j=1}^{q} (n+j)\,|z|^{2n} = q!\frac{1}{(1-|z|^2)^{1+q}}, \\
\langle f_1 f_2 \psi, \psi \rangle &= \bar{z} \sum_{n=0}^{+\infty} \prod_{j=1}^{1+q} (n+j)\,|z|^{2n} = (1+q)!\,\frac{\bar{z}}{(1-|z|^2)^{2+q}}, \\
\langle g_1 g_2 \psi, \psi \rangle &= z \sum_{n=0}^{+\infty} \prod_{j=1}^{1+q} (n+j)\,|z|^{2n} = (1+q)!\,\frac{z}{(1-|z|^2)^{2+q}}, \\
\langle f_1 g_1 \psi, \psi \rangle &= \sum_{n=0}^{+\infty} \prod_{j=0}^{q} (n+j)\,|z|^{2n} = (1+q)!\frac{|z|^2}{(1-|z|^2)^{2+q}}, \\
\langle f_2 g_2 \psi, \psi \rangle &= \sum_{n=0}^{+\infty} \prod_{j=0}^{q} (n+j)\,|z|^{2n} + q\sum_{n=0}^{+\infty} \prod_{j=1}^{q} (n+j)\,|z|^{2n} \\
&= (1+q)!\,\frac{|z|^2}{(1-|z|^2)^{2+q}} + q\,q!\,\frac{1}{(1-|z|^2)^{1+q}}.
\end{aligned}$$

Assim, para $q \geq 0$, os números complexos

$$\frac{\langle B\psi, \psi \rangle}{\langle \psi, \psi \rangle} = (1+q)\,\frac{(c+d)|z|^2 + k\,\bar{z} + l\,z}{1-|z|^2} + qd, \qquad z \in \mathbb{C}_{-1},$$

pertencem a $W(B|_{\Gamma^{(q)}})$. Se $q < 0$, a demonstração é análoga. ■

Começamos por descrever o contradomínio numérico de operadores de emparelhamento auto-adjuntos, restritos a $\Gamma^{(0)}$.

Teorema 2.3.4 *Se o operador de emparelhamento bosónico B definido em* (2.19) *é auto-adjunto e*

$$\lambda_0 = -\frac{1}{2}(c+d) + \frac{1}{2}\,\zeta_{c+d}\sqrt{\Delta_B},$$

então $W(B|_{\Gamma^{(0)}})$ é:

- **a**) $[\lambda_0, +\infty)$, se $\Delta_B > 0$ e $c+d > 0$;
- **b**) $(-\infty, \lambda_0]$, se $\Delta_B > 0$ e $c+d < 0$;
- **c**) $(\lambda_0, +\infty)$, se $\Delta_B = 0$ e $c+d > 0$;
- **d**) $(-\infty, \lambda_0)$, se $\Delta_B = 0$ e $c+d < 0$;
- **e**) $\{0\}$, se $\Delta_B = c+d = 0$;
- **f**) $\mathbb{R}$, se $\Delta_B < 0$.

Demonstração: Como o operador de emparelhamento B é auto-adjunto, pela Proposição 2.3.4, tem-se $c+d \in \mathbb{R}$ e $l = \bar{k}$, logo $\Delta_B \in \mathbb{R}$. Além disso, $W(B|_{\Gamma^{(0)}}) \subseteq \mathbb{R}$ e, sendo o contradomínio numérico convexo, então $W(B|_{\Gamma^{(0)}})$ é um intervalo real, digamos $\mathcal{I}$ (eventualmente degenerado a um ponto ou a toda a recta). Pretendemos caracterizar, caso existam, os pontos extremos de $\mathcal{I}$. Atendendo ao Teorema 1.1.3, se $\lambda \in \mathcal{I}$ é um ponto extremo de $\mathcal{I}$, então λ é um valor próprio de B.

a) Se $\Delta_B > 0$ e $c+d > 0$, pelo Teorema 2.3.1, o valor próprio mínimo do operador de emparelhamento auto-adjunto B, restrito a $\Gamma^{(0)}$, é λ_0 e não existe valor próprio máximo. Portanto, $\{\lambda_0 + iy : \ y \in \mathbb{R}\}$ é a única recta de suporte vertical de $\mathcal{I}$ e tem-se $\mathcal{I} = [\lambda_0, +\infty)$.

b) Se $\Delta_B > 0$ e $c+d < 0$, a prova prossegue de modo análogo à alínea a).

c) Se $\Delta_B = 0$ e $c+d > 0$, então $c+d = 2|k|$ e, mediante cálculos simples, mostra-se que

$$2B \; = \; (c-d)\,(f_1 g_1 - f_2 g_2) + (c+d)\,(f_2+g_1)^*(f_2+g_1) - (c+d)\,\iota^*. \tag{2.45}$$

Quando B se restringe ao subespaço $\Gamma^{(0)}$, a primeira parcela em (2.45) anula-se. Como $c+d > 0$, então $(B - \lambda_0 \iota^*)|_{\Gamma^{(0)}}$ é um operador auto-adjunto semi-definido positivo. Mostra--se que $W((B-\lambda_0\iota^*)|_{\Gamma^{(0)}}) = (0, +\infty)$, ou equivalentemente, $W(C\,|_{\Gamma^{(0)}}) = (0,+\infty)$, onde $C = (f_2+g_1)^*(f_2+g_1)$. Efectivamente, dados

$$w_N \; = \; \sum_{n=1}^{N} \frac{u_n}{n!} f_1^n f_2^n(1) \in \Gamma^{(0)}$$

e $u_0 = u_{N+1} = 0$, temos

$$\frac{\langle C w_N, w_N \rangle}{\langle w_N, w_N \rangle} \; = \; \frac{\sum_{n=0}^{N} (n+1)|u_n + u_{n+1}|^2}{\sum_{n=1}^{N} |u_n|^2} \geq 0,$$

e existem elementos de $W(C\,|_{\Gamma^{(0)}})$ tão próximos de 0 quanto desejarmos. De facto, se $u_n = (-1)^n (N-n)$, $n = 1, \ldots, N$, então

$$\lim_{N\to+\infty} \frac{\langle C w_N, w_N \rangle}{\langle w_N w_N \rangle} \; = \; \lim_{N\to+\infty} \frac{1+2+\cdots+(N+1)}{1+4+\cdots+(N-1)^2} \; = \; 0.$$

Suponhamos que $0 \in W(C|_{\Gamma^{(0)}})$. Nesse caso, zero seria um ponto anguloso de $W(C|_{\Gamma^{(0)}})$ e, pelo Teorema 1.1.3, um valor próprio de C. Isto é, existiria $u \in \Gamma^{(0)} \setminus \{0\}$ tal que $Cu = 0$. Portanto, $\langle Cu, u \rangle = \langle (f_2+g_1)\,u, (f_2+g_1)\,u \rangle$ seria nulo e o mesmo aconteceria com $(f_2+g_1)\,u$, o que contradiz a Proposição 2.3.7. Em virtude de $0 \notin W(C\,|_{\Gamma^{(0)}})$, concluimos que $W(B|_{\Gamma^{(0)}}) = (\lambda_0, +\infty)$.

d) Se $\Delta_B = 0$ e $c+d < 0$, a prova prossegue de modo análogo à alínea c).

e) Se $\Delta_B = c+d = 0$, então $k = 0$ e $B|_{\Gamma^{(0)}} = 0$. Assim, o seu contradomínio numérico é o conjunto singular $\{0\}$.

f) Se $\Delta_B < 0$, como B é auto-adjunto, pelo Lema 2.3.1 tem-se

$$W \; = \; \left\{ \frac{(c+d)|z|^2 + k\bar{z} + \bar{k}z}{1-|z|^2} \; : \; z \in \mathbb{C}_{-1} \right\} \; \subseteq \; W(B|_{\Gamma^{(0)}}) \; \subseteq \; \mathbb{R}.$$

Considerando $r = (1+|z|^2)/(1-|z|^2)$ e $\phi = \arg z - \arg k$, facilmente se verifica que

$$W \;=\; \left\{ \frac{c+d}{2}(r-1) + |k|\sqrt{r^2-1}\,\cos\phi \;:\; 0 \leq \phi < 2\pi,\; r \geq 1 \right\} \;=\; \mathbb{R}.$$

Donde, $W(B|_{\Gamma^{(0)}}) = \mathbb{R}$. ■

Importa ainda realçar que as alíneas a) - f) do Teorema 2.3.4 descrevem o contradomínio numérico das matrizes tridiagonais infinitas Hermíticas

$$\begin{bmatrix} 0 & \bar{k} & 0 & 0 & \dots \\ k & c+d & 2\bar{k} & 0 & \dots \\ 0 & 2k & 2(c+d) & 3\bar{k} & \dots \\ 0 & 0 & 3k & 3(c+d) & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}, \qquad c, d \in \mathbb{R}, \quad k \in \mathbb{C},$$

já que estas são as representações matriciais, na base canónica $\mathcal{B}^0$, dos operadores de emparelhamento auto-adjuntos (2.19) restritos ao subespaço $\Gamma^{(0)}$.

Generalizamos, de seguida, o Teorema 2.3.4 a operadores de emparelhamento B não necessariamente auto-adjuntos, mas ainda restritos ao subespaço $\Gamma^{(0)}$. Como constataremos, existe uma interessante relação entre $W(B|_{\Gamma^{(0)}})$ e o contradomínio numérico-J positivo, $J = \mathrm{diag}\,(1,-1)$, de uma determinada matriz de ordem dois. Tal relação revela-se no papel fundamental que o Teorema do Contradomínio Hiperbólico desempenha na demonstração do próximo teorema. A identidade que aí se obtém entre dois conjuntos de natureza um pouco diferente, o primeiro o contradomínio numérico clássico de um operador ilimitado e o segundo uma das componentes convexas do contradomínio numérico num espaço de Krein de uma matriz de ordem reduzida, conquanto seja surpreendente, é de uma enorme utilidade.

Teorema 2.3.5 *Seja B o operador de emparelhamento bosónico em* (2.19) *e*

$$\lambda_0^\eta \;=\; -\frac{1}{2}(c+d) + \frac{1}{2}\eta\sqrt{\Delta_B}, \qquad \eta \in \{-1, 0, 1\}.$$

Denote-se por l_0 a recta definida pelos pontos λ_0^1 e λ_0^{-1}, excepto o segmento de recta aberto que os une. Sejam

$$\begin{aligned} 2M &= |\Delta_B| - |c+d|^2 + 2|k|^2 + 2|l|^2, \\ 2N &= |\Delta_B| + |c+d|^2 - 2|k|^2 - 2|l|^2. \end{aligned} \tag{2.46}$$

a) *Se $M > 0$ e $N > 0$, então $W(B|_{\Gamma^{(0)}})$ é limitado por um ramo da hipérbole de focos em λ_0^1 e λ_0^{-1}, eixos transverso e não-transverso de comprimento $\sqrt{N}$ e $\sqrt{M}$, respectivamente.*

b) *Se $M > 0$ e $N = 0$, então $W(B|_{\Gamma(0)})$ é:*
 i. *a recta r_0, se $|k| = |l|$;*
 ii. *um semi-plano aberto definido pela recta r_0, se $|k| \neq |l|$;*

em que r_0 é a recta que passa por λ_0^0 e é perpendicular ao segmento que une λ_0^1 a λ_0^{-1}.

c) *Se $M > 0$ e $N < 0$, então $W(B|_{\Gamma(0)})$ é todo o plano complexo.*

d) *Se $M = 0$ e $N > 0$, então $W(B|_{\Gamma(0)})$ é uma semi-recta fechada de extremo λ_0^1 ou λ_0^{-1}, que é uma das componentes disjuntas de l_0.*

e) *Se $M = N = 0$, então $W(B|_{\Gamma(0)})$ é:*
 i. *o conjunto singular $\{0\}$, se $c + d = 0$;*
 ii. *a semi-recta aberta de extremo λ_0^0 contendo a origem, se $c + d \neq 0$.*

Demonstração: Pelo Lema 2.3.1, com $q = 0$, W definido em (2.44) é um subconjunto de $W(B|_{\Gamma(0)})$. Considerando $J = \operatorname{diag}(1, -1)$ e

$$A \;=\; J \begin{bmatrix} 0 & l \\ k & c+d \end{bmatrix},$$

facilmente se verifica que

$$W \;=\; \left\{ \frac{1}{1-|z|^2}\,[1\ \bar{z}]\, JA\, [1\ z]^T :\ z \in \mathbb{C}_{-1} \right\} \;=\; W_J^+(A).$$

Portanto, pode descrever-se W a partir do Teorema do Contradomínio Hiperbólico, tendo em conta a Nota 1.2.2. Seja $P_B = 2|k|^2 + 2|l|^2 - |c+d|^2$. Os valores próprios da matriz A são λ_0^1 a λ_0^{-1}, e temos

$$\begin{aligned} M &= |\lambda_0^1|^2 + |\lambda_0^{-1}|^2 - \operatorname{Tr}(A^{[*]}A) &&= \tfrac{1}{2}\,(|\Delta_B| + P_B)\,, \\ N &= \operatorname{Tr}(A^{[*]}A) \;-\; 2\operatorname{Re}(\overline{\lambda_0^1}\,\lambda_0^{-1}) &&= \tfrac{1}{2}\,(|\Delta_B| - P_B). \end{aligned}$$

Claramente, $M \geq 0$ e

$$|\Delta_B|^2 \;=\; |c+d|^4 + 16|k|^2|l|^2 - 8|k||l||c+d|^2\cos(2\alpha - 2\beta), \tag{2.47}$$

onde $2\alpha = \arg(kl)$ e $\beta = \arg(c+d)$. Pelo Teorema do Contradomínio Hiperbólico, o subconjunto W de $W(B|_{\Gamma(0)})$ é limitado por um ramo de uma hipérbole (possivelmente degenerada). Os casos seguintes podem ocorrer:

1.º caso: $M > 0$ e $N > 0$. Provaremos que $W(B|_{\Gamma(0)}) = W$. Pela Proposição 1.1.1, os vectores próprios unitários associados a um valor próprio extremo de $\operatorname{Re}(e^{i\theta}B)$, $\theta \in [0, 2\pi)$, originam os pontos da fronteira do contradomínio numérico do B. Ora,

$$\operatorname{Re}(e^{i\theta}B) \;=\; c_\theta\, f_1 g_1 + d_\theta\, f_2 g_2 + k_\theta\, f_1 f_2 + \bar{k}_\theta\, g_1 g_2,$$

onde

$$c_\theta = \operatorname{Re}(\mathrm{e}^{i\theta}c), \quad d_\theta = \operatorname{Re}(\mathrm{e}^{i\theta}d) \quad \text{e} \quad 2k_\theta = (k+\bar{l})\cos\theta + i\,(k-\bar{l})\sin\theta.$$

Além disso, $c_\theta + d_\theta = |c+d|\cos(\beta+\theta)$. Tomando $\Delta_\theta = (c_\theta + d_\theta)^2 - 4|k_\theta|^2$, mediante alguns cálculos, obtém-se $2\Delta_\theta = |\Delta_B|\cos(2\theta+\psi) - P_B$, onde $\tan\psi = \mathrm{Im}\Delta_B/\mathrm{Re}\Delta_B$. Consequentemente, $-M \leq \Delta_\theta \leq N$, qualquer que seja $\theta \in [0, 2\pi)$. Seja θ tal que $\Delta_\theta > 0$. Se $c_\theta + d_\theta > 0$, pelo Teorema 2.3.1, o valor próprio mínimo do operador de emparelhamento auto-adjunto $\operatorname{Re}(\mathrm{e}^{i\theta}B)$ é $\lambda_0^\theta = -\frac{1}{2}(c_\theta + d_\theta) + \frac{1}{2}\sqrt{\Delta_\theta}$. Os vectores próprios associados a λ_0^θ são dados por $v_0^\theta = c_0\,\mathrm{e}^{z_\theta f_1 f_2}(1)$, com $c_0 \in \mathbb{C} \setminus \{0\}$, $z_\theta = 0$, se $k_\theta = 0$, e $z_\theta = \lambda_0^\theta/\bar{k}_\theta$, se $k_\theta \neq 0$. Então $z_\theta \in \mathbb{C}_{-1}$ e, seguindo um raciocínio análogo ao utilizado na prova do Lema 2.3.1, para $q = 0$, tem-se

$$\frac{\langle Bv_0^\theta, v_0^\theta\rangle}{\langle v_0^\theta, v_0^\theta\rangle} = \frac{(c+d)|z_\theta|^2 + k\bar{z}_\theta + lz_\theta}{1 - |z_\theta|^2}.$$

Este ponto pertence à fronteira de $W(B|_{\Gamma(0)})$ e pertence a W. À medida que θ varia de 0 a 2π, vão-se descrevendo todos os pontos da fronteira de $W(B|_{\Gamma(0)})$ que pertencem simultaneamente a W. Se $c_\theta + d_\theta < 0$, procede-se a discussão de modo análogo. Assim, $W(B|_{\Gamma(0)}) = W$ é limitado por um ramo da hipérbole de focos λ_0^1 e λ_0^{-1}, eixo transverso de comprimento $\sqrt{N}$ e eixo não-transverso de comprimento $\sqrt{M}$.

2.º caso: $M > 0$ e $N = 0$. Uma vez que $N = 0$, então $M = |\Delta_B| = P_B$. Portanto, $2\Delta_\theta = M(\cos(2\theta+\psi) - 1)$ e facilmente se verifica a existência de $\theta' = -\psi/2 \in [0, 2\pi)$ tal que a função sinusoidal real $f(\theta) = \Delta_\theta$ satisfaz $f(\theta) < 0$, para $\theta \neq \theta'$ e $f(\theta') = 0$. Neste caso, existe uma única recta de suporte de $\overline{W}_B$, mais concretamente a recta r_0. Se $|k| \neq |l|$, então W é um semi-plano aberto definido por esta recta r_0. Pelo Teorema 2.3.4 c) ou d), a fronteira do semi-plano não pertence a $W(B|_{\Gamma(0)})$ e assim $W(B|_{\Gamma(0)})$ coincide com W. Se $|k| = |l| \neq 0$, então W é a recta r_0. Neste caso, Δ_θ e $c_\theta + d_\theta$ anulam-se apenas na direcção $\theta = \pi/2 - \beta \pmod{\pi}$. Pelo Teorema 2.3.4 e), conclui-se que $W(B|_{\Gamma(0)})$ coincide com W. Se $k = l = 0$, então $M = 0$, contradizendo a hipótese.

3.º caso: $M > 0$ e $N < 0$. Quando $N < 0$, não existe nenhuma recta de suporte de W que, por isso, se reduz a todo o plano complexo. Daí, $W(B|_{\Gamma(0)}) = \mathbb{C}$.

4.º caso: $M = 0$ e $N > 0$. Como $M = 0$, tem-se $N = |\Delta_B| = -P_B > 0$. Agora, existe um número infinito de rectas de suporte de W e o ramo da hipérbole, dado pelo Teorema do Contradomínio Hiperbólico, degenera numa semi-recta fechada de extremo λ_0^1 ou λ_0^{-1}, que é uma das componentes disjuntas de l_0. Para $\theta \in [0, 2\pi)$, $2\Delta_\theta = N(\cos(2\theta+\psi) + 1) \geq 0$. Usando argumentos análogos aos da prova do 2.º caso, conclui-se que $W(B|_{\Gamma(0)}) = W$.

5.º caso: $M = 0$ e $N = 0$. Facilmente se vê que $N = \Delta_B$ e, mediante cálculos elementares, tem-se $|k| = |l| = \frac{1}{2}|c+d|$. Se $k = 0$, atendendo ao Teorema 2.3.4 e), concluimos que $W(B) = \{0\}$. Se $k \neq 0$, então W é uma semi-recta aberta contendo a origem e de extremo λ_0^0. A partir de (2.47), vem $0 = |\Delta_B|^2 = 32|k|^4(1 - \cos(2\alpha - 2\beta))$ e daí $\alpha = \beta$. Além disso, $\Delta_\theta = 0$, para $\theta \in [0, 2\pi)$, e $c_\theta + d_\theta$ anula-se apenas na direcção $\theta = \frac{\pi}{2} - \alpha \pmod{\pi}$. Argumentos semelhantes aos utilizados anteriormente permitem mostrar que $W(B|_{\Gamma^{(0)}}) = W$.

6.º caso: $M = 0$ e $N < 0$. Dadas estas hipóteses, tem-se $0 = -M \leq \Delta_\theta \leq N < 0$, o que é impossível. ∎

Traduzindo o Teorema 2.3.5 em termos matriciais, é claro que as suas alíneas apresentam a caracterização do contradomínio numérico das matrizes tridiagonais infinitas

$$\begin{bmatrix} 0 & l & 0 & 0 & \cdots \\ k & c+d & 2l & 0 & \cdots \\ 0 & 2k & 2(c+d) & 3l & \cdots \\ 0 & 0 & 3k & 3(c+d) & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}, \qquad c, d, k, l \in \mathbb{C}.$$

Recorda-se que o Teorema 2.3.4 descreve o contradomínio numérico de operadores de emparelhamento auto-adjuntos restritos a $\Gamma^{(q)}$, no caso particular em que $q = 0$. Em geral, para $q \in \mathbb{Z}$, temos o resultado seguinte.

Teorema 2.3.6 *Seja B o operador de emparelhamento bosónico em* (2.19) *e*

$$\lambda_{0q} = \frac{1+|q|}{2}\left(-(c+d) + \zeta_{c+d}\sqrt{\Delta_B}\right) + \tau_q c + \kappa_q d, \qquad q \in \mathbb{Z}.$$

Se B é auto-adjunto, então $W(B|_{\Gamma^{(q)}})$ é:

a) $[\lambda_{0q}, +\infty)$, se $\Delta_B > 0$ e $c + d > 0$; **b)** $(-\infty, \lambda_{0q}]$, se $\Delta_B > 0$ e $c + d < 0$;
c) $(\lambda_{0q}, +\infty)$, se $\Delta_B = 0$ e $c + d > 0$; **d)** $(-\infty, \lambda_{0q})$, se $\Delta_B = 0$ e $c + d < 0$;
e) $\{\lambda_{0q}\}$, se $\Delta_B = c + d = 0$; **f)** $\mathbb{R}$, se $\Delta_B < 0$.

Demonstração: Seguem-se passos análogos aos da prova do Teorema 2.3.4, recorrendo ao Teorema 2.3.2 em vez do Teorema 2.3.1. ∎

Finalmente, enunciamos o resultado mais geral desta secção. Usando o Lema 2.3.1, o Teorema 2.3.2 e argumentos análogos aos da demonstração do Teorema 2.3.5, caracterizamos o contradomínio numérico do operador de emparelhamento restrito a $\Gamma^{(q)}$, $q \in \mathbb{Z}$. Será possível ainda revelar a existência de uma relação de homotetia entre estes conjuntos.

Teorema 2.3.7 *Seja B o operador de emparelhamento bosónico em* (2.19) *e*

$$\lambda_q^\eta \;=\; \frac{1+|q|}{2}\left(-(c+d)+\eta\sqrt{\Delta_B}\right)+\tau_q c+\kappa_q d, \qquad \eta\in\{-1,0,1\}, \quad q\in\mathbb{Z}.$$

Denote-se por l_q a recta definida pelos pontos λ_q^1 e λ_q^{-1}, excepto o segmento de recta aberto que os une. Sejam

$$\begin{aligned} 2M &= |\Delta_B|-|c+d|^2+2|k|^2+2|l|^2, \\ 2N &= |\Delta_B|+|c+d|^2-2|k|^2-2|l|^2. \end{aligned} \tag{2.48}$$

a) *Se $M>0$ e $N>0$, então $W(B|_{\Gamma^{(q)}})$ é limitado por um ramo da hipérbole de focos em λ_q^1 e λ_q^{-1}, eixos transverso e não-transverso de comprimento $(1+|q|)\sqrt{N}$ e $(1+|q|)\sqrt{M}$, respectivamente.*

b) *Se $M>0$ e $N=0$, então $W(B\,|_{\Gamma^{(q)}})$ é:*

- **i.** *a recta r_q, se $|k|=|l|$;*
- **ii.** *um semi-plano aberto definido pela recta r_q, se $|k|\neq|l|$;*

em que r_q é a recta que passa por λ_q^0 e é perpendicular ao segmento que une λ_q^1 a λ_q^{-1}.

c) *Se $M>0$ e $N<0$, então $W(B\,|_{\Gamma^{(q)}})$ é todo o plano complexo.*

d) *Se $M=0$ e $N>0$, então $W(B\,|_{\Gamma^{(q)}})$ é uma semi-recta fechada de extremo λ_q^1 ou λ_q^{-1}, que é uma das componentes disjuntas de l_q.*

e) *Se $M=N=0$, então $W(B|_{\Gamma^{(0)}})$ é:*

- **i.** *o conjunto singular $\{\lambda_q^0\}$, se $c+d=0$;*
- **ii.** *a semi-recta aberta de extremo λ_q^0 contendo o ponto $\tau_q c+\kappa_q d$, se $c+d\neq 0$.*

Demonstração: Provaremos que

$$W(B\,|_{\Gamma^{(q)}}) \;=\; (1+|q|)\,W(B|_{\Gamma^{(0)}})+\tau_q c+\kappa_q d, \qquad q\in\mathbb{Z}. \tag{2.49}$$

Em primeiro lugar, pelo Lema 2.3.1, o contradomínio numérico do operador de emparelhamento B restrito a $\Gamma^{(q)}$ contém $(1+|q|)\,W+\tau_q c+\kappa_q d$, em que W é o conjunto definido por (2.19). Por outro lado, da demonstração do Teorema 2.3.5, conclui-se que $W=W(B|_{\Gamma^{(0)}})$. Daqui se infere a validade da inclusão:

$$(1+|q|)\,W(B|_{\Gamma^{(0)}})+\tau_q c+\kappa_q d \;\subseteq\; W(B\,|_{\Gamma^{(q)}}), \qquad q\in\mathbb{Z}.$$

De mesmo modo que na demonstração do Teorema 2.3.5, consideramos

$$\mathrm{Re}(\mathrm{e}^{i\theta}B) \;=\; c_\theta\, f_1g_1+d_\theta\, f_2g_2+k_\theta\, f_1f_2+\bar{k}_\theta\, g_1g_2,$$

com $c_\theta=\mathrm{Re}(\mathrm{e}^{i\theta}c)$, $d_\theta=\mathrm{Re}(\mathrm{e}^{i\theta}d)$ e $2k_\theta=(k+\bar{l})\cos\theta+i\,(k-\bar{l})\sin\theta$.

a) Seja $\theta \in [0, 2\pi)$ tal que $\Delta_\theta = (c_\theta + d_\theta)^2 - 4|k_\theta|^2 > 0$. Se $c_\theta + d_\theta > 0$, pelo Teorema 2.3.2, o valor próprio mínimo do operador de emparelhamento auto-adjunto $\mathrm{Re}(\mathrm{e}^{i\theta}B)$ restrito a $\Gamma^{(q)}$, $q \geq 0$, é

$$\lambda_{0q}^\theta \;=\; \frac{q}{2}(d_\theta - c_\theta) - \frac{1}{2}(c_\theta + d_\theta) + \frac{1+q}{2}\sqrt{\Delta_\theta} \;=\; (1+q)\lambda_{00}^\theta + qd_\theta,$$

e os vectores próprios de $\mathrm{Re}(\mathrm{e}^{i\theta}B)$ associados ao valor próprio λ_{0q}^θ são os vectores dados por $v_{0q}^\theta = c_0\,\tilde{f}_2^{\,q}\mathrm{e}^{z_\theta f_1 f_2}(1)$, com $c_0 \in \mathbb{C} \setminus \{0\}$, $z_\theta = 0$, se $k_\theta = 0$, $z_\theta = \lambda_{0q}^\theta/\bar{k}_\theta$, se $k_\theta \neq 0$, e $\tilde{f}_2 = \frac{1}{\sqrt{1-|z_\theta|^2}}(f_2 - \bar{z}_\theta g_1)$. Usando argumentos análogos aos da prova do Lema 2.3.1, temos

$$w_q^\theta \;=\; \frac{\langle Bv_{0q}^\theta, v_{0q}^\theta\rangle}{\langle v_{0q}^\theta, v_{0q}^\theta\rangle} \;=\; (1+q)\frac{(c+d)|z_\theta|^{\,2} + k\,\bar{z}_\theta + l\,z_\theta}{1 - |z_\theta|^{\,2}} + qd, \tag{2.50}$$

que é um ponto da fronteira de $W(B|_{\Gamma^{(q)}})$, $q \geq 0$. Se $c_\theta + d_\theta < 0$, procede-se analogamente. De (2.50), vem a seguinte relação entre os pontos w_q^θ da fronteira de $W(B|_{\Gamma^{(q)}})$, $q > 0$, e os pontos w_0^θ da fronteira de $W(B|_{\Gamma^{(0)}})$:

$$w_q^\theta \;=\; (1+q)w_0^\theta + qd.$$

Isto significa que a curva geradora de fronteira de $W(B|_{\Gamma^{(q)}})$, $q > 0$, se obtém da curva geradora de fronteira de $W(B|_{\Gamma^{(0)}})$ por uma dilatação de razão $1+q$, seguida de uma translação associada a qd. Assim, se mostra a igualdade em (2.49), para $q \geq 0$, e se conclui que $W(B|_{\Gamma^{(q)}})$, $q \geq 0$, é limitado por um ramo da hipérbole de focos λ_q^1 e λ_q^{-1}, e eixos transverso e não-transverso de comprimento $(1+q)\sqrt{N}$ e $(1+q)\sqrt{M}$, respectivamente.

b) Se $|k| \neq |l|$, então $(1+q)\,W + qd$ é um semi-plano aberto definido pela recta r_q. Argumentos análogos aos utilizados na demonstração do Teorema 2.3.6 b) permitem mostrar que a fronteira deste semi-plano não pertence a $W(B|_{\Gamma^{(q)}})$. Assim, $W(B|_{\Gamma^{(q)}})$ coincide com $(1+q)\,W + qd$, para $q \geq 0$. Se $|k| = |l| \neq 0$, então $(1+q)\,W + dq$ é a recta r_q. Neste caso, $\Delta_\theta = (c_\theta + d_\theta)^2 - 4|k_\theta|^2$ e $c_\theta + d_\theta$ anulam-se numa única direcção, donde se pode deduzir a ocorrência de igualdade em (2.49), $q \geq 0$.

c) Dado que $W = \mathbb{C}$, é evidente que $W(B\,|_{\Gamma^{(q)}}) = \mathbb{C}$.

d) Neste caso, o conjunto $(1+q)\,W + qd$ degenera numa das componentes disjuntas do conjunto l_q, que é uma semi-recta fechada de extremo λ_q^1 ou λ_q^{-1}. Já que $\Delta_\theta \geq 0$, para $\theta \in [0, 2\pi)$, argumentos análogos aos usados anteriormente permitem mostrar a igualdade em (2.49), $q \geq 0$.

e) Agora, tal como na prova do 5.º caso do Teorema 2.3.5, tem-se $|k| = |l| = \frac{1}{2}|c+d|$. Se $k = 0$, então $W(B|_{\Gamma^{(q)}}) = \{qd\}$. Se $k \neq 0$, então $(1+q)W + qd$ é a semi-recta aberta de extremo λ_q^0, contendo o ponto qd, e pode concluir-se que $W(B|_{\Gamma^{(q)}}) = (1+q)W + qd$.

Se $q < 0$, a demonstração segue passos similares. ■

Terminada a demonstração do Teorema 2.3.6, importa salientar a relação (2.49) aí obtida, pois esta prova a existência de uma relação de homotetia entre os contradomínios numéricos do mesmo operador de emparelhamento bosónico, quando restrito a diferentes subespaços $\Gamma^{(q)}$, $q \in \mathbb{Z}$. Efectivamente, estes conjuntos são limitados por hipérboles homotéticas, eventualmente degeneradas.

Exemplo 2.3.1 Considere-se o operador de emparelhamento de bosões definido por

$$B \;=\; -if_1g_1 + (2+i)f_2g_2 + f_1f_2 - g_1g_2. \tag{2.51}$$

Neste caso, $\Delta_B = 8$ e $M = N = 4$. Pelo Teorema 2.3.7 a), o conjunto $W(B|_{\Gamma^{(q)}})$, $q \in \mathbb{Z}$, é limitado por um ramo da hipérbole de focos em

$$\lambda_q^{\pm 1} \;=\; -1 \pm \sqrt{2} + q + |q|(i \pm \sqrt{2}),$$

e eixos transverso e não-transverso de igual comprimento $2(1+|q|)$.

Figura 2.1: $\partial W(B|_{\Gamma^{(q)}})$, para B definido em (2.51) e $q = 0, 1, 2, 3, 4$.

Figura 2.2: $\partial W(B|_{\Gamma^{(q)}})$, para B definido em (2.51) e $q = 0, -1, -2, -3, -4$.

Antes de finalizar esta secção, recorda-se a representação matricial, na base canónica $\mathcal{B}^q$, do operador de emparelhamento B restrito a $\Gamma^{(q)}$, $q \in \mathbb{Z}$. Sendo o operador bosónico B definido por (2.19), tal representação é dada pela matriz tridiagonal infinita T_B^q apresentada em (2.20), satisfazendo as condições (2.21). Deste modo, as várias alíneas do Teorema 2.3.7 descrevem minuciosamente o contradomínio numérico das matrizes tridiagonais T_B^q, $q \in \mathbb{Z}$, quaisquer que sejam $c, d, k, l \in \mathbb{C}$.

A título ilustrativo, observa-se que a descrição hiperbólica da fronteira de $W(B|_{\Gamma^{(q)}})$, $q \in \mathbb{Z}$, em que B é o operador definido no Exemplo 2.3.1 por (2.51), caracteriza justamente o contradomínio numérico das matrizes tridiagonais infinitas:

$$T_B^q = \begin{bmatrix} qi & \sqrt{1-q} & 0 & 0 & \cdots \\ -\sqrt{1-q} & 2+qi & \sqrt{2(2-q)} & 0 & \cdots \\ 0 & -\sqrt{2(2-q)} & 4+qi & \sqrt{3(3-q)} & \cdots \\ 0 & 0 & -\sqrt{3(3-q)} & 6+qi & \ddots \\ \vdots & \vdots & \vdots & \ddots & \ddots \end{bmatrix}, \qquad q \leq 0,$$

e $T_B^q = T_B^{-q} + 2q(\,1+i)\,I$, se $q > 0$, onde I denota a matriz identidade infinita. Em particular, o contradomínio numérico da matriz tridiagonal

$$T_B^0 = \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots \\ -1 & 2 & 2 & 0 & \cdots \\ 0 & -2 & 4 & 3 & \cdots \\ 0 & 0 & -3 & 6 & \ddots \\ \vdots & \vdots & \vdots & \ddots & \ddots \end{bmatrix}$$

é limitado por um ramo da hipérbole de focos em $-1 \pm \sqrt{2}$ e de eixos transverso e não-transverso de comprimento dois.

2.3.3 Contradomínio Numérico no Caso Fermiónico

A caracterização efectuada para o contradomínio numérico dos operadores de emparelhamento bosónicos baseou-se nas propriedades espectrais destes operadores anteriormente investigadas. Embora seja possível, para os operadores de emparelhamento fermiónicos, seguir passos semelhantes aos apresentados na subsecção precedente, optamos por estudar o seu contradomínio numérico por um via mais rápida e simples. Basta, para tal, atender à representação matricial do operador na base canónica $\{1, e_1, e_2, e_1 \wedge e_2\}$ da álgebra de Grassmann $\Gamma^{\wedge}$ sobre $\mathbb{C}^2$.

Teorema 2.3.8 *Seja* $B : \Gamma^\wedge \to \Gamma^\wedge$ *o operador de emparelhamento*

$$B \;=\; cf_1g_1 + df_2g_2 + kf_1f_2 + lg_1g_2, \qquad c, d, k, l \in \mathbb{C}, \tag{2.52}$$

e $\Delta_B = (c+d)^2 - 4kl$. *Então* $W(B)$ *é o invólucro convexo dos pontos* c, d *e da elipse* $\mathcal{E}_B$ *(possivelmente degenerada) de focos* $\frac{1}{2}(c+d) \pm \frac{1}{2}\sqrt{\Delta_B}$ *e de eixos maior e menor de comprimento* $\sqrt{M}$ *e* $\sqrt{N}$*, respectivamente, onde*

$$\begin{aligned} 2M &= |c+d|^2 + 2|k|^2 + 2|l|^2 + |\Delta_B|, \\ 2N &= |c+d|^2 + 2|k|^2 + 2|l|^2 - |\Delta_B|. \end{aligned} \tag{2.53}$$

Demonstração: A representação matricial do operador de emparelhamento B na base canónica $\{1, e_1, e_2, e_1 \wedge e_2\}$ da álgebra de Grassmann $\Gamma^\wedge$ é dada pela matriz

$$T_B = \begin{bmatrix} 0 & 0 & 0 & -l \\ 0 & c & 0 & 0 \\ 0 & 0 & d & 0 \\ k & 0 & 0 & c+d \end{bmatrix}, \tag{2.54}$$

que é unitariamente equivalente à soma directa $B_1 \oplus B_2$, onde

$$B_1 = \begin{bmatrix} c & 0 \\ 0 & d \end{bmatrix} \qquad \text{e} \qquad B_2 = \begin{bmatrix} c+d & k \\ -l & 0 \end{bmatrix}.$$

Pela invariância do contradomínio numérico clássico perante transformações de semelhança unitária e pela propriedade W12, relativa à soma directa de matrizes, verifica-se que $W(B) = \mathrm{co}\,(W(B_1) \cup W(B_2))$. Ora, $\mathrm{Tr}(B_2^* B_2) = |c+d|^2 + |k|^2 + |l|^2$ e, dados os valores próprios β_1 e β_2 da matriz B_2, têm-se

$$2|\beta_1^2| + 2|\beta_2|^2 = |\Delta_B| + |c+d|^2 \qquad \text{e} \qquad 4\mathrm{Re}(\beta_1\bar{\beta}_2) = -|\Delta_B| + |c+d|^2.$$

Pelo Teorema do Contradomínio Elíptico, $W(B_2)$ é um disco elíptico (possivelmente degenerado), com os valores próprios de B_2 por focos, eixos maior e menor de comprimento $\sqrt{M}$ e $\sqrt{N}$, respectivamente. Assim, $W(B)$ é o invólucro convexo dos pontos c, d e da elipse $\mathcal{E}_B$ que constitui a fronteira do conjunto $W(B_2)$. ■

Em particular, o operador de emparelhamento fermiónico B definido por (2.52) admite um contradomínio numérico elíptico, cuja fronteira se reduz à elipse $\mathcal{E}_B$ descrita no Teorema 2.3.8, se os pontos c e d pertencem ao domínio delimitado pela curva $\mathcal{E}_B$.

2.4 Primeira Derivação de Matrizes 2×2

Nesta secção, apresenta-se uma nova classe de operadores com contradomínios numéricos elípticos e que podem ser definidos em termos de operadores de criação e destruição de bosões.

Dado um operador linear A em $\mathcal{H}$, o *operador induzido* por A em $\mathcal{H}_{(m)}$ é o operador linear $P_m(A)$ que actua sobre os tensores decomponíveis $x_1 * \cdots * x_m \in \mathcal{H}_{(m)}$ de acordo com a fórmula:

$$P_m(A)(x_1 * \cdots * x_m) \ = \ (Ax_1) * \cdots * (Ax_m).$$

Pela multilineariedade do produto simétrico, tem-se

$$P_m(I_n + tA)(x_1 * \cdots * x_m) \ = \ \sum_{r=0}^{m} t^r \sum_{\alpha \in Q_{r,m}} x_1 * \cdots * (Ax_{\alpha_1}) * \cdots * (Ax_{\alpha_r}) * \cdots * x_m,$$

onde se recorda que $Q_{r,m}$ é o conjunto das sucessões estritamente decrescentes de comprimento r de números inteiros de 1 a m. Pela propriedade universal do produto simétrico, garante-se a existência de um único operador linear $P_m^{(r)}A$ em $\mathcal{H}_{(m)}$, $r = 0, \ldots, m$, tal que

$$P_m^{(r)}A(x_1 * \cdots * x_m) \ = \ \sum_{\alpha \in Q_{r,m}} x_1 * \cdots * (Ax_{\alpha_1}) * \cdots * (Ax_{\alpha_r}) * \cdots * x_m.$$

O operador $P_m^{(r)}A$ diz-se a *r-ésima derivação* de A em $\mathcal{H}_{(m)}$. Neste contexto, a r-ésima derivação de A é completamente determinada, em termos do operador induzido, pela fórmula:

$$P_m(I_n + tA) \ = \sum_{r=0}^{m} t^r P_m^{(r)}A.$$

Se $r = 0$, então $P_m^{(r)}A$ é a aplicação identidade em $\mathcal{H}_{(m)}$.

Em Física, o conceito de derivação é mais relevante do que o conceito de operador induzido e a primeira derivação é especialmente importante. Observáveis de natureza aditiva, tais como a energia cinética, momento linear ou momento angular, ocorrem com frequência. Se A representar uma tal quantidade, para um sistema com uma só partícula, então $P_m^{(1)}A$ representa a quantidade correspondente para um sistema com m partículas, satisfazendo a estatística de Bose-Einstein [18]. Na verdade, se $r = 1$, atendendo a que $Q_{1,m} = \{1, \ldots, m\}$, a *primeira derivação* do operador A em $\mathcal{H}_{(m)}$ satisfaz

$$P_m^{(1)}A(x_1 * \cdots * x_m) \ = \sum_{i=1}^{m} x_1 * \cdots * (Ax_i) * \cdots * x_m.$$

Define-se a primeira derivação de uma matriz $A \in M_n$ em $\mathbb{C}^n_{(m)}$ de modo análogo.

Considere-se, agora, o operador linear definido na álgebra simétrica sobre $\mathbb{C}^2$ por

$$a_{11}f_1g_1 + a_{12}f_1g_2 + a_{21}f_2g_1 + a_{22}f_2g_2, \qquad a_{ij} \in \mathbb{C}, \quad i,j = 1,2, \tag{2.55}$$

à custa de operadores de criação e destruição de bosões. Quando restrito ao subespaço $\mathbb{C}^2_{(m)}$, o operador apresentado em (2.55) é justamente a primeira derivação da matriz de ordem dois de elemento genérico a_{ij} em $\mathbb{C}^2_{(m)}$. O conjunto

$$\mathcal{B} \; = \; \left\{ \frac{e_1^{m-k} * e_2^k}{\sqrt{(m-k)!k!}} : \; k = 0, \dots, m \right\}$$

forma uma base do espaço simétrico $\mathbb{C}^2_{(m)}$. A representação matricial da primeira derivação de uma matriz arbitrária $A = [a_{ij}] \in M_2$ em $\mathbb{C}^2_{(m)}$, na base canónica $\mathcal{B}$, é uma matriz tridiagonal de ordem $(m+1)$ do tipo

$$\begin{bmatrix} a_0 & b_1 & 0 & \dots & 0 \\ c_1 & a_1 & b_2 & \dots & 0 \\ 0 & c_2 & a_2 & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & a_m \end{bmatrix},$$

onde

$$\begin{aligned} a_k &= a_{11}\, m + (a_{22} - a_{11})k, & k &= 0, \dots, m, \\ b_k &= a_{12}\sqrt{k(m-k+1)}, & k &= 1, \dots, m, \\ c_k &= a_{21}\sqrt{k(m-k+1)}, & k &= 1, \dots, m. \end{aligned}$$

2.4.1 Contradomínio Numérico da Primeira Derivação

Com vista a simplificar o estudo do contradomínio numérico da primeira derivação de uma matriz arbitrária de ordem dois, reduz-se este operador a uma forma mais simples. Neste âmbito, intervêm os seguintes operadores definidos à custa de operadores de criação e destruição de bosões. Dado $z \in \mathbb{C}$, sejam $\check{f}_i$, $i = 1,2$, os operadores lineares definidos em $\mathbb{C}^2_{(m)}$ por

$$\check{f}_1 = \frac{1}{\sqrt{1+|z|^2}}\,(f_1 - zf_2) \qquad \text{e} \qquad \check{f}_2 = \frac{1}{\sqrt{1+|z|^2}}\,(f_2 + \overline{z}f_1). \tag{2.56}$$

Os seus operadores adjuntos são dados por

$$\check{g}_1 = \frac{1}{\sqrt{1+|z|^2}}\,(g_1 - \bar{z}g_2) \qquad \text{e} \qquad \check{g}_2 = \frac{1}{\sqrt{1+|z|^2}}\,(g_2 + zg_1), \tag{2.57}$$

respectivamente.

Proposição 2.4.1 *A primeira derivação da matriz Hermítica*

$$A = \begin{bmatrix} a & \bar{b} \\ b & c \end{bmatrix}, \qquad a, c \in \mathbb{R},\ b \in \mathbb{C}, \tag{2.58}$$

em $\mathbb{C}^2_{(m)}$*, transforma-se, por meio de* (2.56) *e* (2.57)*,* $z \in \mathbb{C}$*, na forma*

$$P_m^{(1)} A = (\check{a}\, \check{f}_1 \check{g}_1 + \check{b}\, \check{f}_2 \check{g}_1 + \bar{\check{b}}\, \check{f}_1 \check{g}_2 + \check{c}\, \check{f}_2 \check{g}_2)\,|_{\mathbb{C}^2_{(m)}}, \tag{2.59}$$

onde

$$\begin{aligned} \check{a} &= \frac{1}{1+|z|^2}\,(a + c|z|^2 - b\bar{z} - \bar{b}z), & (2.60)\\ \check{b} &= \frac{1}{1+|z|^2}\,\left((a-c)z + b - \bar{b}z^2\right), & (2.61)\\ \check{c} &= \frac{1}{1+|z|^2}\,(a|z|^2 + c + b\bar{z} + \bar{b}z). & (2.62) \end{aligned}$$

Além disso, $\check{a} + \check{c} = a + c$*.*

Demonstração: A partir de (2.56) e (2.57), obtêm-se as seguintes relações inversas:

$$f_1 = \frac{1}{\sqrt{1+|z|^2}}\,(\check{f}_1 + z\check{f}_2), \qquad f_2 = \frac{1}{\sqrt{1+|z|^2}}\,(\check{f}_2 - \bar{z}\check{f}_1) \tag{2.63}$$

e

$$g_1 = \frac{1}{\sqrt{1+|z|^2}}\,(\check{g}_1 + \bar{z}\check{g}_2), \qquad g_2 = \frac{1}{\sqrt{1+|z|^2}}\,(\check{g}_2 - z\check{g}_1). \tag{2.64}$$

Substituindo (2.63) e (2.64) em $(af_1g_1 + b\, f_2g_1 + \bar{b}\, f_1g_2 + cf_2g_2)\,|_{\mathbb{C}^2_{(m)}}$, tem-se o resultado pretendido. ■

O próximo teorema caracteriza os valores próprios λ_k, $k = 0, \ldots, m$, da primeira derivação de uma matriz Hermítica de ordem dois em $\mathbb{C}^2_{(m)}$. Por conveniência, considera-se a seguinte ordem não crescente: $\lambda_0 \geq \lambda_1 \geq \cdots \geq \lambda_m$.

Teorema 2.4.1 *Os valores próprios da primeira derivação da matriz Hermítica*

$$A = \begin{bmatrix} a & \bar{b} \\ b & c \end{bmatrix}, \qquad a, c \in \mathbb{R},\ b \in \mathbb{C}, \tag{2.65}$$

em $\mathbb{C}^2_{(m)}$*, são dados por*

$$\lambda_k = \frac{m}{2}(a+c) + \frac{m-2k}{2}\sqrt{(a-c)^2 + 4|b|^2}, \qquad k = 0, \ldots, m.$$

Demonstração: Pela Proposição 2.4.1, podemos tomar $P_m^{(1)}A$ na forma (2.59), com $\breve{a}$, $\breve{b}$ e $\breve{c}$ dados por (2.60), (2.61) e (2.62), respectivamente. Facilmente se obtém

$$\breve{a} - \breve{c} = \frac{(a-c)\,(1-|z|^2) - 2b\bar{z} - 2\bar{b}z}{1+|z|^2} \tag{2.66}$$

e é possível determinar um número complexo z tal que $\breve{b}$ se anula. De facto, se $b = 0$ e $a = c$, então $\breve{b} = 0$, para qualquer $z \in \mathbb{C}$. Se $b = 0$ e $a \neq c$, toma-se $z = 0$. Se $b \neq 0$, considera-se

$$z = \frac{a - c \pm \sqrt{(a-c)^2 + 4|b|^2}}{2\bar{b}}.$$

Neste último caso, tem-se $|z|^2 - 1 = \frac{1}{b}(a-c)\,z$ e $|z|^2 + 1 = \pm\frac{1}{b}\sqrt{(a-c)^2+4|b|^2}\,z$. Substituindo z na expressão (2.66), mediante alguns cálculos, vem $\breve{a} - \breve{c} = \mp\sqrt{(a-c)^2+4|b|^2}$. Atendendo a que $\breve{a} + \breve{c} = a + c$, valem as igualdades

$$\breve{a} = \frac{1}{2}(a+c) \pm \frac{1}{2}\sqrt{(a-c)^2+4|b|^2} \qquad \text{e} \qquad \breve{c} = \frac{1}{2}(a+c) \mp \frac{1}{2}\sqrt{(a-c)^2+4|b|^2}.$$

Pode, portanto, reduzir-se a primeira derivação $P_m^{(1)}A$ à forma $(\breve{a}\,\check{f}_1\breve{g}_1 + \breve{c}\,\check{f}_2\breve{g}_2)\,|_{\mathbb{C}^2_{(m)}}$ e os seus valores próprios são do tipo

$$\lambda_k = \breve{a}(m-k) + \breve{c}k, \qquad k = 0, \ldots, m. \tag{2.67}$$

Substituindo $\breve{a}$ e $\breve{c}$ em (2.67), vem

$$\lambda_k = \frac{m}{2}(a+c) \pm \frac{m-2k}{2}\sqrt{(a-c)^2+4|b|^2}, \qquad k = 0, \ldots, m. \tag{2.68}$$

Uma vez que se assume a ordem não crescente para os valores próprios de $P_m^{(1)}A$, substitui-se o sinal '$\pm$' em (2.68) por '$+$'. ■

Prova-se, de seguida, que o contradomínio numérico da primeira derivação de uma matriz arbitrária 2×2 em $\mathbb{C}^2_{(m)}$ é um disco elíptico.

Teorema 2.4.2 *Seja $A \in M_2$ a matriz de valores próprios α_1 e α_2. O contradomínio numérico da primeira derivação de A em $\mathbb{C}^2_{(m)}$ é o disco elíptico (possivelmente degenerado) de focos $m\alpha_1$ e $m\alpha_2$, eixos maior e menor de comprimento*

$$m\sqrt{\mathrm{Tr}\,(A^*A) - 2\,\mathrm{Re}\,(\bar{\alpha}_1\alpha_2)} \qquad e \qquad m\sqrt{\mathrm{Tr}\,(A^*A) - |\alpha_1|^2 - |\alpha_2|^2},$$

respectivamente.

Demonstração: Como já vimos, a primeira derivação da matriz $A = [a_{ij}] \in M_2$ em $\mathbb{C}^2_{(m)}$ é dada pelo operador (2.55). Observa-se que

$$\mathrm{Re}\left(\mathrm{e}^{i\theta} P_m^{(1)} A\right) = (a_\theta\, f_1 g_1 + b_\theta f_1 g_2, + \bar{b}_\theta f_2 g_1 + c_\theta\, f_2 g_2)\,|_{\mathbb{C}^2_{(m)}}, \qquad \theta \in [0, 2\pi),$$

onde $a_\theta = \mathrm{Re}(a_{11}\mathrm{e}^{i\theta})$, $c_\theta = \mathrm{Re}(a_{22}\mathrm{e}^{i\theta})$ e $2b_\theta = a_{12}\mathrm{e}^{i\theta} + a_{21}\,\mathrm{e}^{-i\theta}$. Ora, os valores próprios máximo e mínimo do operador auto-adjunto $\mathrm{Re}\,(\mathrm{e}^{i\theta} P_m^{(1)} A)$ definem as linhas de suporte de $W(P_m^{(1)} A)$. Pelo Teorema 2.4.1, os valores próprios de $\mathrm{Re}\,(\mathrm{e}^{i\theta} P_m^{(1)} A)$ são

$$\lambda_k^\theta = \frac{m}{2}(a_\theta + c_\theta) + \frac{m - 2k}{2}\sqrt{(a_\theta - c_\theta)^2 + 4|b_\theta|^2}, \qquad k = 0, \ldots, m.$$

Portanto, os valores próprios máximo e mínimo do operador $\mathrm{Re}\,(\mathrm{e}^{i\theta} P_m^{(1)} A)$ são $\lambda_0^\theta = mq_\theta$ e $\lambda_m^\theta = -mq_\theta$, respectivamente, onde

$$2q_\theta = (a_\theta + c_\theta) + \sqrt{(a_\theta - c_\theta)^2 + 4|b_\theta|^2}.$$

Isto é, os valores próprios que definem as rectas de suporte de $W(P_m^{(1)} A)$ estão relacionados por uma homotetia de razão m com os valores próprios $\pm\, q_\theta$ que definem as rectas de suporte de $W(P_1^{(1)} A)$. Consequentemente, verifica-se a mesma relação de homotetia entre as curvas geradoras de fronteira destes dois conjuntos e a descrição geométrica de $W(P_m^{(1)} A)$, $m \geq 2$, decorre sem qualquer dificuldade da respectiva descrição no caso $m = 1$. Ora, a representação matricial de $P_1^{(1)} A$, na base canónica de $\mathbb{C}^2$, é obviamente dada pela matriz A, sendo o seu contradomínio numérico descrito pelo Teorema do Contradomínio Elíptico. Conclui-se, assim, que $W(P_m^{(1)} A)$ é um disco elíptico homotético a $W(A)$ e de razão m. ■

Corolário 2.4.1 *Seja A a matriz Hermítica (2.65) e $\Omega_A = (a - c)^2 + 4|b|^2$. O contradomínio numérico da primeira derivação de A em $\mathbb{C}^2_{(m)}$ é:*

(i) $\left[\frac{m}{2}\left(a + c - \sqrt{\Omega_A}\right), \frac{m}{2}\left(a + c + \sqrt{\Omega_A}\right)\right]$, *se* $\Omega_A \neq 0$*;*

(ii) $\left\{\frac{m}{2}(a + c)\right\}$, *se* $\Omega_A = 0$.

Demonstração: Basta notar que $P_m^{(1)} A$ é um operador auto-adjunto, logo $W(P_m^{(1)} A)$ é um intervalo real $[\lambda_m, \lambda_0]$, cujos extremos são os valores próprios mínimo e máximo de $P_m^{(1)} A$. Determinados os valores próprios no Teorema 2.4.1, tem-se de imediato (i). Se $\Omega_A = 0$, então $W(P_m^{(1)} A)$ reduz-se ao conjunto singular em (ii). ■

2.4.2 Contradomínio Numérico-c da Primeira Derivação

E. Brown e I. Spitkovsky [36] mostraram, muito recentemente, que o contradomínio numérico clássico de matrizes tridiagonais é invariante por troca de duas entradas simétricas relativamente à diagonal principal. Antes disso, Mao-Ting Chien e Hiroshi Nakazato [43] provaram um resultado análogo para o contradomínio numérico-c, $c \in \mathbb{R}^n$, de matrizes tridiagonais com diagonal principal de zeros. Generalizamos este resultado a matrizes tridiagonais arbitrárias, qualquer que seja $c \in \mathbb{R}^n$. Para o efeito, utilizamos o lema seguinte, sobre a representação paramétrica da fronteira de $W_c(A)$, $c \in \mathbb{R}^n$ [40, 42].

Lema 2.4.1 [42] *Seja* $c = (c_1, \ldots, c_n) \in \mathbb{R}^n$, *tal que* $c_1 \geq \cdots \geq c_n$, $A \in M_n$ *e*

$$\lambda(\theta) = \sum_{i=1}^{n} c_i \lambda_i(\theta),$$

onde $\lambda_1(\theta) \geq \cdots \geq \lambda_n(\theta)$ *são os valores próprios de* $\mathrm{Re}(\mathrm{e}^{i\theta} A)$, $\theta \in [0, 2\pi)$. *Então* $W_c(A)$ *é o invólucro convexo da curva* $\{X(\theta) + i\,Y(\theta) : \theta \in [0, 2\pi)\}$, *onde*

$$\begin{aligned} X(\theta) &= \cos\theta\,\lambda(\theta) - \sin\theta\,\lambda'(\theta), \\ Y(\theta) &= -\sin\theta\,\lambda(\theta) - \cos\theta\,\lambda'(\theta). \end{aligned}$$

A curva $\{X(\theta) + i\,Y(\theta) : \theta \in [0, 2\pi)\}$ definida no Lema 2.4.1 denomina-se a *curva geradora da fronteira* de $W_c(A)$, $c \in \mathbb{R}^n$.

Teorema 2.4.3 *Se* A *é uma matriz tridiagonal do tipo*

$$\begin{bmatrix} a_1 & b_1 & 0 & \ldots & 0 \\ d_1 & a_2 & b_2 & \ldots & 0 \\ 0 & d_2 & a_3 & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & a_n \end{bmatrix},$$

então o contradomínio numérico-c, $c \in \mathbb{R}^n$ *é invariante por troca das entradas* b_j *e* d_j, *para cada* $j = 1, \ldots, n-1$.

Demonstração: Para cada $j = 1, \ldots, n-1$, seja

$$\widehat{A}_j = \begin{bmatrix} a_1 & b_1 & 0 & 0 & 0 & 0 \\ d_1 & \ddots & \ddots & 0 & 0 & 0 \\ 0 & \ddots & a_j & d_j & 0 & 0 \\ 0 & 0 & b_j & a_{j+1} & \ddots & 0 \\ 0 & 0 & 0 & \ddots & \ddots & b_{n-1} \\ 0 & 0 & 0 & 0 & d_{n-1} & a_n \end{bmatrix} \qquad (2.69)$$

a matriz tridiagonal que apenas difere de A por troca das entradas b_j e d_j. Por indução sobre n, facilmente se verifica que o polinómio característico de $\mathrm{Re}\,(\mathrm{e}^{i\theta} A)$ coincide com o de $\mathrm{Re}\,(\mathrm{e}^{i\theta}\widehat{A}_j)$, $\theta \in [0, 2\pi)$. Se $n = 2$, a afirmação é claramente verdadeira. Suponhamos a afirmação válida para todas as matrizes tridiagonais de ordem não superior a n. Prova-se que também vale para as de ordem $n+1$, isto é, para as matrizes

$$B = \begin{bmatrix} A & b \\ d & a_{n+1} \end{bmatrix},$$

em que b e d denotam os vectores coluna e linha, com todas as entradas nulas, excepto possivelmente a última igual a $b_n \in \mathbb{C}$ ou a $d_n \in \mathbb{C}$, respectivamente. Então

$$\widehat{B}_j = \begin{bmatrix} \widehat{A}_j & b \\ d & a_{n+1} \end{bmatrix}, \quad j = 1, \ldots, n-1, \qquad \text{e} \qquad \widehat{B}_n = \begin{bmatrix} A & d^T \\ b^T & a_{n+1} \end{bmatrix}$$

diferem de B por troca das entradas b_j e d_j, $j = 1, \ldots, n$, respectivamente. Denotemos por $A' = A(n)$ a matriz que se obtém de A, eliminando a linha e coluna n. O polinómio característico de $\mathrm{Re}\,(\mathrm{e}^{i\theta}B)$ é dado por

$$p_{n+1,B}(t) = \big(t - \mathrm{Re}\,(a_{n+1}\mathrm{e}^{i\theta})\big)\; p_{n,A}(t) - \frac{1}{4}\left|b_n\mathrm{e}^{i\theta} + \bar{d}_n\mathrm{e}^{-i\theta}\right|^2 p_{n-1,A'}(t), \tag{2.70}$$

em que $p_{n,A}(t) = \det\big(tI_n - \mathrm{Re}(\mathrm{e}^{i\theta}A)\big)$ e $p_{n-1,A'}(t) = \det\big(tI_{n-1} - \mathrm{Re}(\mathrm{e}^{i\theta}A')\big)$. Pela hipótese indutiva, tem-se

$$p_{n,A}(t) = p_{n,\widehat{A}_j}(t) \qquad \text{e} \qquad p_{n-1,A'}(t) = p_{n-1,\widehat{A}'_j}(t).$$

Logo, $p_{n+1,B}(t)$ coincide com o polinómio característico de $\mathrm{Re}\,(\mathrm{e}^{i\theta}\widehat{B}_j)$, $j = 1, \ldots, n-1$. Por outro lado, como nenhuma diferença ocorre, se b_n e d_n inverterem posições em (2.70), então $p_{n+1,B}(t)$ também é o polinómio característico de $\mathrm{Re}\,(\mathrm{e}^{i\theta}\widehat{B}_n)$.

Concluída a demonstração por indução, verifica-se que coincidem os valores próprios de $\mathrm{Re}\,(\mathrm{e}^{i\theta}A)$ e os de $\mathrm{Re}\,(\mathrm{e}^{i\theta}\widehat{A}_j)$, qualquer que seja $\theta \in [0, 2\pi)$. Usando o Lema 2.4.1, tem-se $W_c(A) = W_c(\widehat{A}_j)$, $c \in \mathbb{R}^n$. ■

Agora, caracteriza-se o contradomínio numérico-c da primeira derivação de uma matriz arbitrária de ordem dois em $\mathbb{C}^2_{(m)}$, com excepção do caso em que são idênticas as duas primeiras componentes de $c \in \mathbb{R}^{m+1}$.

Teorema 2.4.4 *Sejam $c = (c_0, \ldots, c_m) \in \mathbb{R}^{m+1}$, com $c_0 \neq c_1$, $s_c = \sum_{k=0}^{m}(m-2k)\,c_k$ e $A \in M_2$ de valores próprios α_1 e α_2. Então $W_c\,(P_m^{(1)}A)$ é o disco elíptico (possivelmente degenerado) de focos*

$$s_c\,\alpha_1 + \mathrm{Tr}(A)\sum_{k=0}^{m} kc_k \qquad e \qquad s_c\,\alpha_2 + \mathrm{Tr}(A)\sum_{k=0}^{m} kc_k,$$

e de eixos maior e menor de comprimento

$$|s_c|\sqrt{\mathrm{Tr}\,(A^*A) - 2\,\mathrm{Re}(\bar{\alpha}_1\alpha_2)} \qquad e \qquad |s_c|\sqrt{\mathrm{Tr}\,(A^*A) - |\alpha_1|^2 - |\alpha_2|^2}, \tag{2.71}$$

respectivamente.

Demonstração: Dada $A = [a_{ij}]$, centra-se a atenção na matriz $B = A - \frac{1}{2}\,\mathrm{Tr}(A)\,I_2$, pois a multilinearidade do produto simétrico garante que $P_m^{(1)}A = P_m^{(1)}B + \frac{1}{2}\,m\,\mathrm{Tr}(A)\,\mathrm{id}$, onde id denota a aplicação identidade em $\mathbb{C}^2_{(m)}$. Verifica-se que

$$\mathrm{Re}\left(\mathrm{e}^{i\theta}P_m^{(1)}B\right) \;=\; \left(a_\theta(f_1g_1 - f_2g_2) + b_\theta f_1g_2 + \bar{b}_\theta f_2g_1\right)\Big|_{\mathbb{C}^2_{(m)}}, \qquad \theta \in [0, 2\pi),$$

onde $a_\theta = \mathrm{Re}\left((a_{11} - a_{22})\,\mathrm{e}^{i\theta}\right)$ e $2b_\theta = a_{12}\mathrm{e}^{i\theta} + a_{21}\mathrm{e}^{-i\theta}$. Pelo Teorema 2.4.1, os valores próprios da primeira derivação da matriz Hermítica $\mathrm{Re}\left(\mathrm{e}^{i\theta}B\right)$, restrita a $\mathbb{C}^2_{(m)}$, são da forma

$$\beta_k^\theta \;=\; \frac{m-2k}{2}\,\sqrt{2a_\theta + 4|b_\theta|^2}, \qquad k = 0, \ldots, m.$$

Pelo Lema 2.4.1, $W_c(P_m^{(1)}B)$ é o invólucro convexo de $\{X_m(\theta) + i\,Y_m(\theta) : \theta \in [0, 2\pi)\}$, onde

$$X_m(\theta) \;=\; \cos\theta\,\lambda_m(\theta) - \sin\theta\,\lambda_m'(\theta), \tag{2.72}$$

$$Y_m(\theta) \;=\; -\sin\theta\,\lambda_m(\theta) - \cos\theta\,\lambda_m'(\theta), \tag{2.73}$$

com

$$\lambda_m(\theta) \;=\; \sum_{k=0}^{m} c_k\,\beta_k^\theta \;=\; \frac{s_c}{2}\sqrt{2a_\theta + 4|b_\theta|^2}.$$

Ora, $(c_0 - c_1)\,\lambda_m(\theta) = s_c\,\lambda_1(\theta)$. Sendo $c_0 \neq c_1$, de (2.72) e (2.73), derivam-se as seguintes relações:

$$X_m(\theta) = \frac{s_c}{c_0 - c_1}X_1(\theta) \qquad \text{e} \qquad Y_m(\theta) = \frac{s_c}{c_0 - c_1}Y_1(\theta),$$

isto é, ocorre uma relação de homotetia de razão $|s_c|/|c_0 - c_1|$ entre as curvas geradoras de fronteira dos conjuntos $W_c\,(P_m^{(1)}B)$, $m \geq 2$, e $W_{c'}(P_1^{(1)}B) = W_{c'}(B)$, $c' = (c_0, c_1)$. Sejam $\pm\beta_1$ os valores próprios de B. Pelo Teorema 1.3.1, o conjunto $W_{c'}(B)$ é o disco elíptico (possivelmente degenerado) de centro na origem, com focos nos pontos-σ $\pm(c_0 - c_1)\beta_1$, eixos maior e menor de comprimento

$$|c_0 - c_1|\sqrt{\mathrm{Tr}\,(B^*B) + 2|\beta_1|^2} \qquad \text{e} \qquad |c_0 - c_1|\sqrt{\mathrm{Tr}\,(B^*B) - 2|\beta_1|^2},$$

respectivamente. Portanto, $\partial W_c\,(P_m^{(1)}B)$ é uma elipse de focos $\pm\,s_c\,\beta_1$. Como os valores próprios de A são da forma $\frac{1}{2}\mathrm{Tr}(A) \pm \beta_1$, facilmente se verificam as identidades:

$$\mathrm{Tr}\,(A^*A) - |\alpha_1|^2 - |\alpha_2|^2 \;=\; \mathrm{Tr}\,(B^*B) - 2|\beta_1|^2,$$

$$\mathrm{Tr}\,(A^*A) - 2\,\mathrm{Re}\,(\bar{\alpha}_1\alpha_2) \;=\; \mathrm{Tr}\,(B^*B) + 2|\beta_1|^2$$

e, assim, (2.71) determina o comprimento dos eixos maior e menor da fronteira elíptica de $W_c\,(P_m^{(1)}B)$. A partir da relação existente entre $P_m^{(1)}A$ e $P_m^{(1)}B$, tem-se

$$W_c\big(P_m^{(1)}A\big) \;=\; W_c\big(P_m^{(1)}B\big) \;+\; \mathrm{Tr}(A)\,\frac{m}{2}\sum_{k=0}^{m} c_k,$$

donde se conclui o pretendido. ■

Estamos, agora, em condições de apresentar uma nova classe de matrizes tridiagonais com contradomínio numérico-c elíptico, $c \in \mathbb{R}^n$, como consequência quase imediata dos Teoremas 2.4.3 e 2.4.4.

Corolário 2.4.2 *Sejam $c = (c_0, \ldots, c_m) \in \mathbb{R}^{m+1}$, com $c_0 \neq c_1$, $s_c = \sum_{k=0}^{m}(m-2k)\,c_k$ e $A = [a_{ij}] \in M_2$ de valores próprios α_1 e α_2. Dada uma partição de $\{1, \ldots, m\}$ em conjuntos τ e τ', o contradomínio numérico-c da matriz tridiagonal*

$$D_\tau = \begin{bmatrix} a_0 & b_1 & 0 & \ldots & 0 \\ d_1 & a_1 & b_2 & \ldots & 0 \\ 0 & d_2 & a_2 & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & a_m \end{bmatrix},$$

onde

$$\begin{array}{lll} a_k = a_{11}\,m + (a_{22}-a_{11})k, & k = 0, \ldots, m, & \\ b_k = a_{12}\sqrt{k(m-k+1)}, & k \in \tau', & b_k = a_{21}\sqrt{k(m-k+1)}, \quad k \in \tau, \\ d_k = a_{21}\sqrt{k(m-k+1)}, & k \in \tau', & d_k = a_{12}\sqrt{k(m-k+1)}, \quad k \in \tau, \end{array}$$

é o disco elíptico (possivelmente degenerado) de focos

$$s_c\,\alpha_1 + \mathrm{Tr}(A)\sum_{k=0}^{m} kc_k \qquad e \qquad s_c\,\alpha_2 + \mathrm{Tr}(A)\sum_{k=0}^{m} kc_k,$$

e de eixos maior e menor de comprimento

$$|s_c|\sqrt{\mathrm{Tr}\,(A^*A) - 2\,\mathrm{Re}(\bar{\alpha}_1\alpha_2)} \qquad e \qquad |s_c|\sqrt{\mathrm{Tr}\,(A^*A) - |\alpha_1|^2 - |\alpha_2|^2}, \tag{2.74}$$

respectivamente.

Demonstração: A representação matricial da primeira derivação da matriz A, restrita a $\mathbb{C}^2_{(m)}$, na base canónica $\mathcal{B}$, é dada pela matriz tridiagonal $D_\emptyset$. Pelo Teorema 2.4.3, o seu contradomínio numérico-c, especificado no Teorema 2.4.4, é invariante por troca das entradas b_k e c_k, $k \in \tau$, $\tau \subseteq \{1, \ldots, m\}$. Ora, D_τ é a matriz tridiagonal que se obtém de $D_\emptyset$ por troca das entradas das diagonais secundárias b_k e d_k, $k \in \tau$. ■

Capítulo 3

Desigualdades Matriciais em Física

No one really understands entropy. Therefore, if you know what you mean by it and you use it when you are in an argument, you will win every time.

John von Neumann

3.1 Entropia

O conceito de entropia[1] foi introduzido em termodinâmica por Rudolf Clausius, em 1865, e alguns dos maiores passos na consolidação deste conceito foram dados por Ludwig Boltzmann e J. W. Gibbs [112]. Desde então, a entropia extravasou a mecânica estatística, foi reformulada e generalizada, e invadiu áreas tão distintas como a teoria da informação, os sistemas dinâmicos, a teoria ergódica, a biologia, a economia, as ciências sociais e humanas.

Em mecânica quântica, os estados puros de um sistema físico são descritos por vectores de um espaço de Hilbert, sendo os estados mistos descritos por matrizes semi-definidas positivas de traço um. Tais matrizes chamam-se *matrizes densidade* e os seus valores próprios são as probabilidades de o sistema em causa se encontrar nos estados puros descritos pelos correspondentes vectores próprios.

Em sistemas clássicos discretos ou comutativos, podem identificar-se os estados físicos com matrizes de densidade diagonais que estão associadas a vectores de probabilidade p, isto é,

$$p \ = \ (p_1, \ldots, p_n), \qquad p_i \ \geq \ 0, \quad i = 1, \ldots, n, \qquad \sum_{i=1}^{n} p_i \ = \ 1. \tag{3.1}$$

[1]A palavra *entropia* tem origem na palavra grega $\eta\ \tau\rho o\pi\eta$, que significa transformação, e foi intencionalmente escolhida por Clausius dada a sua semelhança com a palavra energia [7].

Neste contexto, define-se a *entropia de Shannon* de p por

$$S(p) \;=\; -\sum_{i=1}^{n} p_i \log p_i, \tag{3.2}$$

adoptando-se a convenção $x \log x = 0$, se $x = 0$. Claude Shannon introduziu esta função, em 1948, no seu trabalho pioneiro em teoria da informação *Uma Teoria Matemática da Comunicação* [120], como a medida de eficiência de um sistema de comunicação. Nesse tratado, onde Shannon desenvolveu, num golpe de génio, a teoria da informação e transmissão de sinais digitais, a abordagem axiomática dada à função (3.2) não tem precedentes.

Em sistemas quânticos, define-se a *entropia* de um estado misto, ou da matriz densidade A que o descreve, por

$$S(A) \;=\; -\mathrm{Tr}\,(A \log A)\,.$$

A noção de entropia quântica foi introduzida, em 1927, por Jonh von Neumann [111], em termodinâmica, para traduzir o grau de desordem de um sistema, muito antes de Shannon recriar a correspondente quantidade clássica. Aliás, *medida de informação* foi o primeiro termo usado por Shannon; só alterado por sugestão do próprio von Neumann, que observou a analogia entre a formulação matemática da função de Shannon e da entropia de um sistema físico[2].

Obviamente, a entropia quântica é invariante por transformações de semelhança unitária, ou seja, $S(U^*AU) = S(A)$, para cada matriz unitária U. Ora, toda a matriz densidade A é unitariamente semelhante a uma matriz diagonal cujos elementos principais são os valores próprios de A. Pode, assim, expressar-se a entropia de $A \in M_n$ em termos dos seus valores próprios λ_i, $i = 1, \ldots, n$, por

$$S(A) \;=\; -\sum_{i=1}^{n} \lambda_i \log \lambda_i.$$

A entropia mínima $S(A) = 0$ ocorre se e só se um dos valores próprios da matriz densidade A for um e os restantes forem nulos, e a entropia máxima $S(A) = \log n$ ocorre se e só se todos os valores próprios de A forem iguais entre si, ou seja, para $A = I_n/n$.

Entropia Relativa

A entropia que relaciona dois estados de um sistema comutativo, descritos por vectores de probabilidade p e q, ambos satisfazendo condições do tipo (3.1), é geralmente

[2] Segundo consta, von Neumann terá, ironicamente, dito a Shannon: *deve chamar-lhe entropia por duas razões: primeiro, porque essa mesma função matemática é já utilizada em termodinâmica, com esse nome; segundo, e mais importante, porque pouca gente sabe realmente o que é a entropia e, se usar esse termo numa discussão, sairá sempre a ganhar.*

conhecida por *entropia relativa de Kullback-Leibler* [88], também chamada *divergência de informação*, e tem a seguinte expressão:

$$S(p,q) \;=\; \sum_{i=1}^{n} p_i \log \frac{p_i}{q_i},$$

em que se convencionam $x \log x = 0$, se $x = 0$, e $x \log y = +\infty$, se $x \neq 0$.

O seu análogo quântico, ou não comutativo, é a *entropia relativa de Umegaki* [129] definida para matrizes densidade A e B por

$$S(A,B) \;=\; \operatorname{Tr}\left(A(\log A - \log B)\right).$$

Sendo A uma matriz densidade em M_n e, portanto, de traço um, verifica-se a igualdade: $S(A) - S(I_n/n) = -S(A, I_n/n)$. Como a condição do traço unitário não é essencial, consideram-se, de agora em diante, as noções de entropia e de entropia relativa no contexto geral das matrizes semi-definidas positivas. Claramente, vale a relação: $S(A) = -S(A, I_n)$.

Outra variante introduzida por Jun Ichi Fujii e Eizaburo Kamei [54], em teoria da informação não comutativa, é a *entropia relativa para operadores* definida por

$$\hat{S}(A|B) \;=\; A^{1/2} \log(A^{-1/2} B A^{-1/2}) A^{1/2}.$$

Previamente, Nakamura e Umegaki [106] consideraram a entropia para operadores, dada pela fórmula $-A \log A$, e Belavkin e Staszewski [28] utilizaram, no contexto de sistemas físicos descrito por álgebras-C^*, a quantidade $-\operatorname{Tr} \hat{S}(A|B)$.

É óbvio que a entropia relativa de Umegaki satisfaz a seguinte propriedade de invariância:

$$S(U^*AU, U^*BU) \;=\; S(A,B),$$

qualquer que seja a matriz unitária U. Além disso, se as matrizes A e B comutam entre si, então são simultaneamente diagonalizáveis e tem-se

$$-\operatorname{Tr} \hat{S}(A|B) \;=\; S(A,B) \;=\; \sum_{i=1}^{n} \lambda_i \log \frac{\lambda_i}{\gamma_i},$$

em que λ_i e γ_i, $i = 1, \ldots, n$, são os valores próprios (com vectores próprios em comum) de A e B, respectivamente.

Média da Potência de Matrizes

Seja $0 \leq \alpha \leq 1$. Define-se a *média da potência-α* das matrizes $A > 0$ e $B \geq 0$ por

$$A \#_\alpha B \;=\; A^{1/2} \left(A^{-1/2} B A^{-1/2}\right)^\alpha A^{1/2},$$

noção que se estende a matrizes $A \geq 0$ e $B \geq 0$, por continuidade, da seguinte forma:

$$A\#_\alpha B \;=\; \lim_{\epsilon \downarrow 0}(A+\epsilon I_n)\#_\alpha B.$$

Na teoria da média de operadores lineares positivos estabelecida por Kubo e Ando [87], a média da potência-α é a que está associada à função monótona $f(x) = x^\alpha$. Para qualquer $0 \leq \alpha \leq 1$, tem-se $A\#_\alpha B \geq 0$ e verifica-se a propriedade $A\#_\alpha B = B\#_{1-\alpha}A$. Em particular, $A\#_0 B = A$, $A\#_1 B = B$ e $A\#B = A\#_{1/2}B$ é a *média geométrica* das matrizes A e B, terminologia justificada por $A\#B = (AB)^{1/2}$, quando $AB = BA$. A média da potência-α verifica ainda a seguinte propriedade:

$$\left.\frac{d}{d\alpha}A\#_\alpha B\right|_{\alpha=0} \;=\; \hat{S}(A|B). \tag{3.3}$$

3.2 Desigualdade de Furuta e Majoração Logarítmica

Com vista a apresentar uma conhecida desigualdade matricial que preserva uma relação de ordem, importante pelas suas aplicações, introduzimos a seguinte notação. Escreve-se $A \geq B$ para significar que $A, B \in M_n$ são matrizes Hermíticas e $A - B$ é semi-definida positiva, ou equivalentemente, $\langle Ax, x\rangle \geq \langle Bx, x\rangle$, $x \in \mathbb{C}^n$. Observa-se que a relação $\geq$ é de ordem parcial no conjunto das matrizes Hermíticas.

Define-se uma classe de funções reais de variável real, cuja extensão às matrizes Hermíticas preserva a anterior relação de ordem; as chamadas *funções matriciais monótonas* são as funções reais contínuas f, definidas num intervalo $\mathcal{I} \subseteq \mathbb{R}$, que satisfazem a propriedade:

$$A \;\geq\; B \quad\Rightarrow\quad f(A) \;\geq\; f(B),$$

em que $f(A)$ se define pelo cálculo funcional usual [32, p.112] para uma matriz Hermítica A de valores próprios em $\mathcal{I}$.

Listam-se alguns factos fundamentais, embora triviais, envolvendo, para $A \in M_n$, a norma do operador $\|A\|$ e o *raio espectral*

$$\rho(A) \;=\; \max\{|\lambda| : \lambda \in \sigma(A)\},$$

coincidentes, quando A é uma matriz Hermítica, e iguais ao maior valor próprio $\lambda_1(A)$, se $A \geq 0$. (Para mais pormenores sobre F1-F4, consultar [32].)

F1. (Desigualdade de Browne, 1928) Se $A \in M_n$, então $\rho(A) \leq \|A\|$.

F2. Dada $A \geq 0$, tem-se $\lambda_1(A) \leq 1$ se e só se $A \leq I_n$.

F3. Dada $A \in M_n$, tem-se $\|A\| \leq 1$ se e só se $AA^* \leq I_n$.

F4. Se $A \geq B$, então $X^*AX \geq X^*BX$, qualquer que seja $X \in M_n$.

Desigualdade de Löwner-Heinz

Sempre que $0 \leq \alpha \leq 1$, a função $f(x) = x^\alpha$, $x \in [0, +\infty)$, é uma função matricial monótona. Este facto foi originalmente obtido, em 1934, por Karl Löwner [99]. Em 1951, Erhard Heinz [76] estendeu-o a operadores lineares em espaços de dimensão infinita. É vulgarmente conhecido por *desigualdade de Löwner-Heinz*. Desde então, surgiram na literatura outras demonstrações baseadas em métodos analíticos ou puramente algébricos (vide, por exemplo, [1, 32, 55, 84]). A prova que apresentamos é adaptada da de Gert K. Pedersen [115].

Lema 3.2.1 (Desigualdade de Löwner-Heinz, 1934) *Se $A \geq B \geq 0$ e $0 \leq \alpha \leq 1$, então $A^\alpha \geq B^\alpha$.*

Demonstração: Seja S o conjunto dos expoentes $\alpha \in [0, 1]$ para os quais $A \geq B \geq 0$ implica $A^\alpha \geq B^\alpha$. Claramente, $0 \in S$ e $1 \in S$. Basta mostrar que S é convexo. Sem perda de generalidade, suponhamos A invertível. Seja $C_\alpha = A^{-\alpha/2}B^{\alpha/2}$. Se $\alpha \in S$, então $A^\alpha \geq B^\alpha$, o que implica $C_\alpha C_\alpha^* \leq I_n$, ou ainda $\|C_\alpha\| \leq 1$. Analogamente, se $\beta \in S$, então $\|C_\beta\| \leq 1$. Pela desigualdade de Browne e pela submultiplicatividade da norma, tem-se

$$\rho(C_\alpha C_\beta^*) \;\leq\; \|C_\alpha C_\beta^*\| \;\leq\; \|C_\alpha\|\|C_\beta^*\| \;\leq\; 1.$$

Além disso, dado $\gamma = (\alpha + \beta)/2$, tem-se $\lambda_1(C_\gamma C_\gamma^*) = \rho(C_\alpha C_\beta^*) \leq 1$, donde $C_\gamma C_\gamma^* \leq I_n$ e, portanto, $A^\gamma \geq B^\gamma$, ou seja, $\gamma \in S$. ■

Resulta, sem dificuldade, da desigualdade de Löwner-Heinz que $(A, B) \mapsto A\#_\alpha B$ é uma função monótona na segunda variável e, pela propriedade $A\#_\alpha B = B\#_{1-\alpha}A$, também na primeira variável.

A título de curiosidade, observa-se que o Lema 3.2.1 é equivalente a resultados bem conhecidos, alguns dos quais envolvendo $A, B \geq 0$ e $0 \leq \alpha \leq 1$, como a *desigualdade de Cordes* [55, 61]:

$$\|A^\alpha B^\alpha\| \;\leq\; \|AB\|^\alpha,$$

a *desigualdade de Heinz-Kato* [55, 84]:

$$T^*T \;\leq\; A^2 \;\wedge\; T^*T \;\leq\; B^2 \quad\Rightarrow\quad \|\langle Tx, y\rangle\| \;\leq\; \|A^\alpha x\|\|B^{1-\alpha}y\|, \quad x, y \in \mathbb{C}^n,$$

e a *desigualdade de Jensen* [74]:

$$(X^*AX)^\alpha \geq X^*AX,$$

válida qualquer que seja $X \in M_n$, satisfazendo $\|X\| \leq 1$.

Muito recentemente, Tsuyoshi Ando [4] estabeleceu a desigualdade de Löwner de tipo indefinido, no âmbito de matrizes Hermíticas-J de valores próprios reais não negativos.

Desigualdade de Furuta

A desigualdade de Löwner-Heinz não é, em geral, válida para $\alpha > 1$. Não obstante, verifica-se que

$$A \geq B \geq 0 \quad \Rightarrow \quad f(A^r) \geq f(B^r), \qquad r > 1, \tag{3.4}$$

onde $f(X) = (AXA)^{1/r}$ ou $f(X) = (BXB)^{1/r}$. Esta resposta afirmativa de Takayuki Furuta [58] à conjectura, para $r = 2$, de Chan e Kwong [37], esteve na base da mais geral *desigualdade de Furuta.* Esta última é uma generalização notável da desigualdade de Löwner-Heinz, tem variadíssimas aplicações e vem publicada na *Encyclopaedia of Mathematics* [50]. Conhecem-se várias demonstrações da desigualdade de Furuta (vide [63, §3.2]), entre as quais uma de uma só página [59] que usa a decomposição polar. Optámos pela prova baseada na média da potência-α [60], por a considerarmos mais elegante.

Teorema 3.2.1 (Desigualdade de Furuta, 1987) *Se $A \geq B \geq 0$, então*

$$\left(B^{s/2}A^rB^{s/2}\right)^\alpha \geq \left(B^{s/2}B^rB^{s/2}\right)^\alpha, \tag{3.5}$$

$$\left(A^{s/2}A^rA^{s/2}\right)^\alpha \geq \left(A^{s/2}B^rA^{s/2}\right)^\alpha, \tag{3.6}$$

para $r, s \geq 0$ e $0 \leq \alpha \leq 1$ tais que $(r+s)\,\alpha \leq 1+s$.

Demonstração [60]**:** Se $0 \leq r \leq 1$, então o resultado é consequência óbvia da desigualdade de Löwner-Heinz e de F4.

Seja $r \geq 1$, $0 \leq s \leq 1$ e $\alpha = (1+s)/(r+s)$. Sem perda de generalidade, suponhamos A e B invertíveis. Pela desigualdade de Löwner-Heinz, $A \geq B \geq 0$ implica $B^{-s} \geq A^{-s}$. A monotonia da média da potência-α na primeira variável assegura que

$$B^{-s}\#_\alpha A^r \geq A^{-s}\#_\alpha A^r = A \geq B. \tag{3.7}$$

Atendendo a F4, tem-se

$$\left(B^{s/2}A^rB^{s/2}\right)^\alpha = B^{s/2}\left(B^{-s}\#_\alpha A^r\right)B^{s/2} \geq B^{1+s} = \left(B^{s/2}B^rB^{s/2}\right)^\alpha, \tag{3.8}$$

o que prova (3.5) para qualquer $0 \leq s \leq 1$. Seja, agora, $t = 2s+1$ e $\alpha_1 = (1+t)/(r+t)$. Então $\alpha_1 = 2\alpha(1+\alpha)$ e

$$\begin{aligned}
\left(B^{t/2}A^rB^{t/2}\right)^{\alpha_1} &= B^{t/2}\left(B^{-t}\#_{\alpha_1}A^r\right)B^{t/2} \\
&= B^{(t-s)/2}\left(B^{-1-s}\#_{\alpha_1}(B^{s/2}A^rB^{s/2})\right)B^{(t-s)/2} \\
&\geq B^{(t-s)/2}\left((B^{s/2}A^rB^{s/2})^{-\alpha}\#_{\alpha_1}(B^{s/2}A^rB^{s/2})\right)B^{(t-s)/2} \qquad (3.9) \\
&= B^{(t-s)/2}\left(B^{s/2}A^rB^{s/2}\right)^{\alpha}B^{(t-s)/2} \\
&= B^{t/2}\left(B^{-s}\#_{\alpha}A^r\right)B^{t/2} \geq B^{1+t} = \left(B^{t/2}B^rB^{t/2}\right)^{\alpha_1}, \qquad (3.10)
\end{aligned}$$

em que a desigualdade (3.9) resulta de tomar inversas em (3.8), da monotonia da média da potência-α_1 na primeira variável e de F4; e (3.10) é imediato de (3.7) por meio de F4. Assim se prova (3.5) com $1 \leq t \leq 3$ no lugar s. Repetindo este procedimento, mostra-se (3.5) para $r \geq 1$, $s \geq 0$ e $\alpha = \tilde{\alpha}$, onde $\tilde{\alpha} = (1+s)/(r+s)$. Daqui, pela desigualdade de Löwner-Heinz, com $0 \leq \alpha/\tilde{\alpha} \leq 1$, conclui-se (3.5) para $r \geq 1$, $s \geq 0$ e $\alpha \leq \tilde{\alpha}$. Finalmente, obtém-se (3.6) ao tomar inversas na desigualdade que resulta de aplicar (3.5) a $B^{-1} \geq A^{-1} > 0$. ■

Pode interpretar-se o Teorema 3.2.1 no sentido de que vale a afirmação (3.4), para

$$f(X) = \left(A^{s/2}XA^{s/2}\right)^{\alpha} \quad \text{ou} \quad f(X) = \left(B^{s/2}XB^{s/2}\right)^{\alpha}, \qquad s \geq 0, \quad 0 \leq \alpha \leq 1,$$

desde que $(r+s)\,\alpha \leq 1+s$. Claramente, obtém-se a desigualdade de Löwner-Heinz da desigualdade de Furuta, quando $r=1$ e $s=0$. Além disso, o domínio definido para r, s, α no Teorema 3.2.1 é o melhor possível [125].

A desigualdade de Furuta é extremamente rica em aplicações. Alguns exemplos são a caracterização da ordem caótica $A \gg B$, isto é, $\log A \geq \log B$; generalização da desigualdade de Heinz-Kato; extensão de desigualdades traciais de Kosaki; desigualdades de tipo Kantorovich; várias desigualdades de normas e de operadores (consultar [63] para mais pormenores). Aplicaremos a desigualdade de Furuta, mais concretamente o corolário seguinte, na prova do último resultado desta secção.

Corolário 3.2.1 *Se $A \geq B \geq 0$, então $A^{1+q} \geq \left(A^{p/2}B^{p/q}A^{p/2}\right)^{q/p}$, $0 < q \leq p$.*

Demonstração: Basta considerar (3.6) com $r = p/q$, $s = p$ e $\alpha = q/p$. ■

Como generalização da sua desigualdade, Furuta [62] obteve também a Grande desigualdade de Furuta, uma fórmula paramétrica interpoladora entre a desigualdade de Furuta e a desigualdade de Ando-Hiai [2, Teorema 3.5].

Majoração de Matrizes

Dada uma matriz Hermítica $H \in M_n$, fixamos a seguinte ordem não crescente para os seus valores próprios: $\lambda_1(H) \geq \cdots \geq \lambda_n(H)$.

Sejam H e K matrizes Hermíticas em M_n. Se

$$\sum_{i=1}^{k} \lambda_i(H) \;\leq\; \sum_{i=1}^{k} \lambda_i(K), \qquad k = 1, \ldots, n, \tag{3.11}$$

e se $\mathrm{Tr}(H) = \mathrm{Tr}(K)$, diz-se que K *majora* H, o que se denota por $H \prec K$. Se (3.11) se verificar, sem que ocorra necessariamente igualdade para $k = n$, diz-se que K *majora fracamente* H e escreve-se $H \prec_w K$.

Majoração Logarítmica de Matrizes

Sejam A e B matrizes semi-definidas positivas em M_n. Se

$$\prod_{i=1}^{k} \lambda_i(A) \;\leq\; \prod_{i=1}^{k} \lambda_i(B), \qquad k = 1, \ldots, n, \tag{3.12}$$

e se $\det(A) = \det(B)$, diz-se que B *majora-log* A e usa-se a notação $A \prec_{(\log)} B$. Justifica--se esta terminologia pelo facto da majoração-log entre matrizes definidas positivas A e B equivaler à majoração usual entre as matrizes Hermíticas $\log A$ e $\log B$.

É bem conhecido que a majoração logarítmica $A \prec_{(\log)} B$ implica a majoração fraca $A \prec_w B$ [79, Proposição 1.3]. Este conceito é a base de uma técnica poderosa para derivar desigualdades $|||A||| \leq |||B|||$ para normas unitariamente invariantes $|||\cdot|||$ e, em particular, desigualdades traciais. Uma norma $|||\cdot|||$ diz-se *unitariamente invariante* em M_n se $|||A||| = |||UAV|||$, para cada $A \in M_n$ e quaisquer que sejam $U, V \in M_n$ unitárias.

Matrizes Compostas

Introduzimos, de seguida, a noção de matriz composta que permite estender resultados envolvendo o maior valor próprio de uma matriz a resultados envolvendo produtos dos seus k maiores valores próprios. Seja $A \in M_n$ e $n_k = \binom{n}{k}$, $k = 1, \ldots, n$. A *k-ésima matriz composta* de A, denotada por $A^{(k)}$, é a matriz $n_k \times n_k$ cujas entradas são dadas por $\det A(\mathcal{R}, \mathcal{S})$, com $\mathcal{R}, \mathcal{S} \subseteq \{1, \ldots, n\}$ conjuntos de índices de cardinalidade k ordenados lexicograficamente [103]. Listamos algumas propriedades básicas das matrizes compostas:

C1. **(Fórmula de Binet-Cauchy, 1812)** Dadas $A, B \in M_n$, $A^{(k)}B^{(k)} = (AB)^{(k)}$, $k = 1, \ldots, n$.

C2. Se $A \geq 0$, então $A^{(k)} \geq 0$ e $\big(A^{(k)}\big)^q = (A^q)^{(k)}$, para cada $q > 0$, $k = 1, \ldots, n$.

C3. Se $A \geq 0$, então $\lambda_1\big(A^{(k)}\big) = \prod_{i=1}^{k} \lambda_i(A)$, $k = 1, \ldots, n$.

Atendendo à propriedade C3, pode reescrever-se (3.12) da definição de majoração logarítmica, em linguagem de matrizes compostas na forma:

$$\lambda_1\big(A^{(k)}\big) \ \leq \ \lambda_1\big(B^{(k)}\big), \qquad k = 1, \ldots, n.$$

Uma Majoração Logarítmica Tipo Ando-Hiai

Por exemplo, Huzihiro Araki [5], ao estender a desigualdade de Lieb-Thirring[3], utilizou a técnica das matrizes compostas e obteve, para $A, B \geq 0$, a seguinte majoração logarítmica:

$$\big(A^{1/2}BA^{1/2}\big)^s \ \prec_{(\log)} \ A^{s/2}B^sA^{s/2}, \qquad s \geq 1, \tag{3.13}$$

equivalente a

$$\big(A^{q/2}B^qA^{q/2}\big)^{1/q} \ \prec_{(\log)} \ \big(A^{p/2}B^pA^{p/2}\big)^{1/p}, \qquad 0 < q \leq p. \tag{3.14}$$

Utilizando técnicas de Tsuyoshi Ando e Fumio Hiai [2], obtém-se de seguida uma majoração logarítmica deste tipo. A sua demonstração apoia-se no Corolário 3.2.1 da desigualdade de Furuta.

Teorema 3.2.2 *Se $A, B \geq 0$, então*

$$A^{(1+q)/2}B^qA^{(1+q)/2} \ \prec_{(\log)} \ A^{1/2}(A^{p/2}B^pA^{p/2})^{q/p}A^{1/2}, \qquad 0 < q \leq p. \tag{3.15}$$

Demonstração: Pretendemos mostrar que

$$\lambda_1\left(A^{(1+q)/2}B^qA^{(1+q)/2}\right) \ \leq \ \lambda_1\left(A^{1/2}(A^{p/2}B^pA^{p/2})^{q/p}A^{1/2}\right), \qquad 0 < q \leq p. \tag{3.16}$$

Nesse caso, basta substituir A e B em (3.16) pelas respectivas k-ésimas matrizes compostas, $k = 1, \ldots, n$, atender às propriedades C1 - C3 e observar que os determinantes de ambos os membros de (3.15) coincidem, para se ter o resultado.

Suponhamos, por absurdo, que existe $\gamma > 0$ tal que

$$\lambda_1\left(A^{(1+q)/2}B^qA^{(1+q)/2}\right) \ > \ \gamma \ \geq \ \lambda_1\left(A^{1/2}(A^{p/2}B^pA^{p/2})^{q/p}A^{1/2}\right). \tag{3.17}$$

[3] $\mathrm{Tr}\,((AB)^s) \leq \mathrm{Tr}\,(A^sB^s)$, $s \geq 1$, obtida por Lieb e Thirring [98], para $A, B \geq 0$, e aplicada para obter desigualdades para os momentos dos valores próprios do Hamiltoniano de Schrödinger.

Podemos tomar A e B invertíveis. Dadas $C = \gamma^{1/(1+q)}A$ e $D = \gamma^{-2/q}B$, dividindo todos os membros de (3.17) por γ, e tendo em conta que

$$\begin{aligned}\lambda_1\left(C^{(1+q)/2}D^qC^{(1+q)/2}\right) &= \gamma^{-1}\,\lambda_1\left(A^{(1+q)/2}B^qA^{(1+q)/2}\right),\\ \lambda_1\left(C^{1/2}(C^{p/2}D^pC^{p/2})^{q/p}C^{1/2}\right) &= \gamma^{-1}\,\lambda_1\left(A^{1/2}(A^{p/2}B^pA^{p/2})^{q/p}A^{1/2}\right),\end{aligned}$$

tem-se

$$\lambda_1\left(C^{(1+q)/2}D^qC^{(1+q)/2}\right) > 1 \geq \lambda_1\left(C^{1/2}(C^{p/2}D^pC^{p/2})^{q/p}C^{1/2}\right). \tag{3.18}$$

De (3.18), vem $I_n \geq C^{1/2}\left(C^{p/2}D^pC^{p/2}\right)^{q/p}C^{1/2}$, o que implica $C^{-1} \geq \left(C^{p/2}D^pC^{p/2}\right)^{q/p}$. Pelo Corolário 3.2.1 da desigualdade de Furuta, tem-se $C^{-1-q} \geq D^q$, ou equivalentemente, $I_n \geq C^{(1+q)/2}D^qC^{(1+q)/2}$. Tal contradiz, atendendo a F2, a primeira desigualdade em (3.18). Em conclusão, vale (3.16). ■

3.3 Desigualdades Traciais Logarítmicas

Apresentam-se, seguidamente, algumas desigualdades traciais logarítmicas que envolvem as noções de entropia relativa introduzidas na Secção 3.1. Começa-se por referir um resultado devido a Fumio Hiai e Dénes Petz [77, Teorema 3.5] e, mais tarde, complementado por Tsuyoshi Ando e Fumio Hiai [2, Teorema 5.1]. Delineamos uma nova prova para este resultado que fornece um limite superior para a entropia relativa de Umegaki, utilizando para esse efeito o Teorema 3.2.2 e o seguinte lema auxiliar.

Lema 3.3.1 *Se $A > 0$ e H é um matriz Hermítica, então* $\mathrm{Tr}(A\,[\log A, H]\,H) \geq 0$.

Demonstração: Suponhamos $A = \mathrm{diag}\,(a_1, \ldots, a_n) > 0$ e $X = [x_{ij}] \in H_n$. Neste caso,

$$\mathrm{Tr}\left(A\log A\,X^2\right) = \sum_{k,j=1}^{n} a_k\log a_k|x_{kj}|^2, \qquad \mathrm{Tr}\left(A\,X\log A\,X\right) = \sum_{k,j=1}^{n} a_k\log a_j|x_{kj}|^2.$$

Como $[\log A, X] = \log A\,X - X\log A$, pela linearidade do traço, temos

$$\begin{aligned}\mathrm{Tr}\left(A\,[\log A, X]\,X\right) &= \sum_{k<j} a_k\left(\log a_k - \log a_j\right)|x_{kj}|^2 + \sum_{j<k} a_k\left(\log a_k - \log a_j\right)|x_{kj}|^2,\\ &= \sum_{k<j}\left(a_k - a_j\right)\left(\log a_k - \log a_j\right)|x_{kj}|^2 \geq 0,\end{aligned}$$

pois $|x_{jk}| = |x_{kj}|$, $j, k = 1, \ldots, n$, e $(a-b)(\log a - \log b) \geq 0$, sempre que $a, b > 0$.

Em geral, existe uma matriz unitária $U \in M_n$ que diagonaliza A. Sejam $D = U^*AU$ e $X = U^*HU$. Então $\mathrm{Tr}\left(A\,[\log A, H]\,H\right) = \mathrm{Tr}\left(D\,[\log D, X]\,X\right) \geq 0$, pois D é diagonal. ■

Teorema 3.3.1 *Se $A \geq 0$, $B \geq 0$, então, para cada $p > 0$, tem-se*

$$\mathrm{Tr}\,(A(\log A + \log B)) \;\leq\; \frac{1}{p}\,\mathrm{Tr}\left(A \log\left(A^{p/2} B^p A^{p/2}\right)\right) \tag{3.19}$$

e o segundo membro de (3.19) *converge decrescentemente para o primeiro membro, à medida que $p \downarrow 0$.*

Demonstração: Podemos supor que $B > 0$ [77, Teorema 3.5] e considerar $A > 0$, dada a continuidade de $\mathrm{Tr}\left(A \log(A^{p/2} B^p A^{p/2})\right)$ em $A \geq 0$. Como a majoração logarítmica implica a majoração fraca, a partir do Teorema 3.2.2, tem-se

$$\mathrm{Tr}\left(A^{1+q} B^q\right) \;\leq\; \mathrm{Tr}\left(A\left(A^{p/2} B^p A^{p/2}\right)^{q/p}\right), \qquad 0 < q \leq p. \tag{3.20}$$

Ora, ambos os membros de (3.20) são iguais a $\mathrm{Tr}(A)$ quando $q = 0$, daí

$$\frac{d}{dq}\mathrm{Tr}\left(A^{1+q} B^q\right)\bigg|_{q=0} \;\leq\; \frac{d}{dq}\mathrm{Tr}\left(A\left(A^{p/2} B^p A^{p/2}\right)^{q/p}\right)\bigg|_{q=0}, \qquad p > 0,$$

e (3.19) é consequência directa do cálculo destas derivadas.

Argumentos padrão permitem estender $\log\left(A^{p/2} B^p A^{p/2}\right)$ a uma função analítica $f(p)$ numa vizinhança da origem. Mediante alguns cálculos, obtém-se o seguinte desenvolvimento em série de potências:

$$f(p) \;=\; p\,(\log A + \log B) \;+\; \frac{p^3}{24}\Big[\,[\log A, \log B]\,, 2\log B + \log A\Big] \;+\; \cdots,$$

para qualquer $p \in \mathbb{R}$ numa vizinhança da origem. (Recorda-se que $[X, Y] = XY - YX$ denota o comutador das matrizes X e Y.) No que se refere ao desenvolvimento em série de potências de $\frac{1}{p}\,\mathrm{Tr}\,(Af(p))$, é nulo o coeficiente de p e, pelo facto de $[A, \log A] = 0$ e pela propriedade cíclica do traço, facilmente se conclui que o coeficiente de p^2 é igual a

$$\frac{1}{24}\,\mathrm{Tr}\left(A\Big[\,[\log A, \log B]\,, 2\log B + \log A\Big]\right) \;=\; \frac{1}{6}\,\mathrm{Tr}\left(A\,[\log A, \log B]\log B\right) \;\geq\; 0,$$

onde o sinal não negativo resulta de aplicar o Lema 3.3.1 a $H = \log B$. Assim, o segundo membro de (3.19) decresce para o primeiro membro, quando $p \downarrow 0$. ■

Fumio Hiai [78] provou que ocorre igualdade na desigualdade tracial logarítmica (3.19) se e só se $AB = BA$.

Usando a terminologia de entropia relativa, pode reescrever-se (3.19) na forma

$$S(A, B) \;\leq\; -\frac{1}{p}\mathrm{Tr}\left(\hat{S}(A^p|B^p)A^{1-p}\right), \qquad p > 0, \tag{3.21}$$

para $A, B > 0$. Se $p = 1$, (3.21) estabelece uma relação entre a entropia relativa de Umegaki e a correspondente versão de Belavkin e Staszewski: $S(A, B) \leq -\mathrm{Tr}\,\hat{S}(A|B)$.

Corolário 3.3.1 *Se $A, B > 0$ e $0 \leq \alpha \leq 1$, então*

$$S\left(A, (A^r \#_\alpha B^s)^{t/p}\right) \leq -\frac{1}{p} \operatorname{Tr}\left(\hat{S}\left(A^p | (A^r \#_\alpha B^s)^t\right) A^{1-p}\right),$$

para $r, s \geq 0$, $t \geq 1$ e $p > 0$.

Demonstração: O Corolário é imediato de (3.21), substituindo B por $(A^r \#_\alpha B^s)^{t/p}$. ■

Recupera-se (3.21) do Corolário 3.3.1, no caso particular $\alpha = 1$, $p = st$ e $s \neq 0$.

Usando o Corolário 3.3.1, a versão paramétrica da *fórmula de Lie-Trotter* [119]:

$$e^{X+Y} = \lim_{p \to 0} \left(e^{pX/2} e^{pY} e^{pX/2}\right)^{1/p}, \qquad X, Y \in M_n,$$

e a sua variante para a média da potência-α [77, Lema 3.3]:

$$e^{(1-\alpha)H + \alpha K} = \lim_{p \to 0} \left(e^{pH} \#_\alpha e^{pK}\right)^{1/p}, \qquad H, K \in H_n, \tag{3.22}$$

apresenta-se uma prova alternativa para mais uma desigualdade tracial logarítmica de Ando e Hiai [2, Teorema 5.3].

Corolário 3.3.2 *Se $A \geq 0$, $B > 0$, $0 \leq \alpha \leq 1$ e $p > 0$, então*

$$\frac{1}{p} \operatorname{Tr}\left(A \log\left(A^p \#_\alpha B^p\right)\right) + \frac{\alpha}{p} \operatorname{Tr}\left(A \log\left(A^{p/2} B^{-p} A^{p/2}\right)\right) \geq \operatorname{Tr}\left(A \log A\right) \tag{3.23}$$

e o primeiro membro de (3.23) *converge para* $\operatorname{Tr}(A \log A)$, *quando $p \downarrow 0$.*

Demonstração: Seja $A > 0$ e considere-se o Corolário 3.3.1 no caso particular $t = 1$ e $p = r = s$. Obtém-se (3.23), visto que

$$\begin{aligned} S(A, (A^p \#_\alpha B^p)^{1/p}) &= \operatorname{Tr}(A \log A) - \frac{1}{p} \operatorname{Tr}\left(A \log\left(A^p \#_\alpha B^p\right)\right), \\ \operatorname{Tr}(\hat{S}\left(A^p | A^p \#_\alpha B^p\right) A^{1-p}) &= -\alpha \operatorname{Tr}\left(A \log\left(A^{p/2} B^{-p} A^{p/2}\right)\right). \end{aligned}$$

Agora, estudamos a convergência quando $p \downarrow 0$. Aplicando a fórmula (3.22) às matrizes Hermíticas $H = \log A$ e $K = \log B$, e atendendo à continuidade da função logaritmo, tem-se

$$\lim_{p \downarrow 0} \log\left(A^p \#_\alpha B^p\right)^{1/p} = (1-\alpha) \log A + \alpha \log B. \tag{3.24}$$

Por outro lado, aplicando a versão paramétrica da fórmula de Lie-Trotter a $X = \log A$ e $Y = -\log B$, vem:

$$\lim_{p \downarrow 0} \log\left(A^{p/2} B^{-p} A^{p/2}\right)^{1/p} = \log A - \log B. \tag{3.25}$$

De (3.24) e (3.25), facilmente se conclui a convergência do primeiro membro de (3.23) para o segundo, quando $p \downarrow 0$. Se $A \geq 0$, por uma perturbação, toma-se $A + \epsilon I_n$, $\epsilon > 0$, e por um argumento de continuidade, conclui-se o pretendido. ■

Em terminologia de entropia relativa, o Corolário 3.3.2 estabelece que

$$S\left(A, (A^p \#_\alpha B^p)^{1/p}\right) \leq -\frac{\alpha}{p} \operatorname{Tr}\left(\hat{S}(A^p|B^p)A^{1-p}\right). \tag{3.26}$$

A prova do Corolário 3.3.2 mostra ainda que ambos os membros de (3.26) convergem para $\alpha\, S(A,B)$ à medida que $p \downarrow 0$. Claramente, basta tomar $\alpha = 1$ em (3.26) para se recuperar (3.21).

3.4 Sobre a Desigualdade de Peierls-Bogoliubov

Um problema importante do ponto de vista da mecânica estatística é o cálculo do valor da *função de partição*:

$$Z = \operatorname{Tr}(\mathrm{e}^{-\hat{H}/\Theta}),$$

onde $\hat{H}$ é o Hamiltoniano do sistema físico, isto é, uma matriz Hermítica, e $\Theta = k_B T$, com k_B a constante de Boltzmann e T a temperatura absoluta. O cálculo do valor da função de partição nem sempre é fácil. Por conveniência, pode tomar-se $\Theta = -1$ e pode revelar-se mais simples determinar a quantidade relacionada $\operatorname{Tr}(\mathrm{e}^H)$, para uma aproximação conveniente H do Hamiltoniano $\hat{H}$, isto é, $\hat{H} = H + K$.

A *desigualdade de Peierls-Bogoliubov* fornece informação útil sobre $\operatorname{Tr}(\mathrm{e}^{H+K})$ a partir de $\operatorname{Tr}(\mathrm{e}^H)$. Esta desigualdade estabelece, para duas matrizes Hermíticas H e K, que

$$\operatorname{Tr}(\mathrm{e}^H) \exp \frac{\operatorname{Tr}(\mathrm{e}^H K)}{\operatorname{Tr}(\mathrm{e}^H)} \leq \operatorname{Tr}(\mathrm{e}^{H+K}). \tag{3.27}$$

Esta desigualdade bem conhecida será estendida no Corolário 3.4.4 do seguinte teorema de Ando e Hiai [3]. Observa-se que a igualdade ocorre em (3.27) se e só se K é uma matriz escalar.

Por conveniência, perante um produto de duas matrizes semi-definidas positivas do tipo AB, por este ser unitariamente semelhante à matriz Hermítica $A^{1/2}BA^{1/2}$, escreve-se $\lambda_i(AB)$ em vez de $\lambda_i(A^{1/2}BA^{1/2})$, $i = 1, \ldots, n$.

Teorema 3.4.1 *Se $A_1, A_2, B_1, B_2 \geq 0$ e $0 \leq \alpha \leq 1$, então*

$$\prod_{i=1}^{k} \lambda_i\left((A_1 \#_\alpha A_2)(B_1 \#_\alpha B_2)\right) \leq \prod_{i=1}^{k} \left(\lambda_i(A_1 B_1)\right)^{1-\alpha} \left(\lambda_i(A_2 B_2)\right)^{\alpha}, \qquad k = 1, \ldots, n, \tag{3.28}$$

ocorrendo igualdade para $k = n$.

Demonstração [3]**:** Se $k = n$, é claro que ambos os membros de (3.28) coincidem com $(\det A_1 \det B_1)^{1-\alpha}(\det A_2 \det B_2)^{\alpha}$. Basta provar que

$$\lambda_1\left((A_1 \#_\alpha A_2)\,(B_1 \#_\alpha B_2)\right) \;\leq\; (\lambda_1(A_1B_1))^{1-\alpha}(\lambda_1(A_2B_2))^{\alpha}, \qquad 0 \leq \alpha \leq 1, \tag{3.29}$$

pois ao substituir A_1, A_2, B_1, B_2 em (3.29) pelas respectivas k-ésimas matrizes compostas, $k = 1, \ldots, n$, atendendo às propriedades C1 - C3, prova-se (3.28).

Sejam $\gamma_1 = \lambda_1(A_1B_1)$ e $\gamma_2 = \lambda_1(A_2B_2)$. Podemos supor A_1 e A_2 invertíveis. Dadas $C_1 = \gamma_1^{-1}A_1$ e $C_2 = \gamma_2^{-1}A_2$, tem-se

$$\lambda_1\left((C_1 \#_\alpha C_2)\,(B_1 \#_\alpha B_2)\right) \;=\; \gamma_1^{\alpha-1}\gamma_2^{-\alpha}\lambda_1((A_1 \#_\alpha A_2)\,(B_1 \#_\alpha B_2)).$$

Portanto, provar (3.29) equivale a provar

$$\lambda_1\left((C_1 \#_\alpha C_2)\,(B_1 \#_\alpha B_2)\right) \;\leq\; 1, \qquad 0 \leq \alpha \leq 1. \tag{3.30}$$

Obviamente, $\lambda_1(C_1B_1)^{1-\alpha} = \lambda_1(C_2B_2)^{\alpha} = 1$, donde resultam $C_1B_1 \leq I_n$ e $C_2B_2 \leq I_n$, ou seja, $B_1 \leq C_1^{-1}$ e $B_2 \leq C_2^{-1}$. Pela monotonia da média da potência-α em ambas as variáveis, conclui-se que $B_1 \#_\alpha B_2 \leq C_1^{-1} \#_\alpha C_2^{-1} = (C_1 \#_\alpha C_2)^{-1}$, o que implica $(C_1 \#_\alpha C_2)\,(B_1 \#_\alpha B_2) \leq I_n$ e, portanto, vale (3.30). ■

A partir do Teorema 3.4.1, Ando e Hiai obtiveram outra desigualdade tracial logarítmica, que apresentamos reescrita no Corolário 3.4.1 de forma condensada, recorrendo à entropia relativa.

Corolário 3.4.1 *Se $A_1, B_1 \geq 0$ e $A_2, B_2 > 0$, então*

$$\mathrm{Tr}\left(\hat{S}(A_1|A_2)\,B_1\right) + \mathrm{Tr}\left(\hat{S}(B_1|B_2)\,A_1\right) \;\leq\; -\frac{1}{n}S\left(\mathrm{Tr}\,(A_1B_1)\,I_n, \mathrm{Tr}\,(A_2B_2)\,I_n\right). \tag{3.31}$$

Demonstração [3]**:** Por continuidade, suponhamos A_1 e B_1 invertíveis. De (3.28), vem

$$\mathrm{Tr}\left((A_1 \#_\alpha A_2)(B_1 \#_\alpha B_2)\right) \leq \left(\mathrm{Tr}\,(A_1B_1)\right)^{1-\alpha}\left(\mathrm{Tr}\,(A_2B_2)\right)^{\alpha}.$$

Para $\alpha = 0$, ambos os membros da anterior desigualdade coincidem com $\mathrm{Tr}(A_1B_1)$, pelo que

$$\left.\frac{d}{d\alpha}\mathrm{Tr}\left((A_1 \#_\alpha A_2)(B_1 \#_\alpha B_2)\right)\right|_{\alpha=0} \;\leq\; \mathrm{Tr}\,(A_1B_1)\,\left.\frac{d}{d\alpha}\left(\frac{\mathrm{Tr}\,(A_2B_2)}{\mathrm{Tr}\,(A_1B_1)}\right)^{\alpha}\right|_{\alpha=0}.$$

Do cálculo destas derivadas, tendo em conta a propriedade (3.3), tem-se o resultado. ■

Corolário 3.4.2 *Se $A_1, B_1 \geq 0$ e $A_2, B_2 > 0$, então*

$$\begin{aligned} \frac{1}{n} S\left(\operatorname{Tr}\left(A_1 B_1\right) I_n, \operatorname{Tr}\left(\left(A_1 \#_\alpha A_2\right)\left(B_1 \#_\beta B_2\right)\right) I_n\right) & \leq -\alpha \operatorname{Tr}\left(\hat{S}(A_1 | A_2)\, B_1\right) \\ & \quad -\beta \operatorname{Tr}\left(\hat{S}(B_1 | B_2)\, A_1\right), \end{aligned}$$

para cada $0 \leq \alpha \leq 1$ e $0 \leq \beta \leq 1$.

Demonstração: Basta substituir as matrizes A_2 e B_2 em (3.31) por $A_1 \#_\alpha A_2$ e $B_1 \#_\beta B_2$, respectivamente, e atender às identidades $\operatorname{Tr}(\hat{S}(A_1 | A_1 \#_\alpha A_2)\, B_1) = \alpha \operatorname{Tr}(\hat{S}(A_1 | A_2)\, B_1)$ e $\operatorname{Tr}(\hat{S}(B_1 | B_1 \#_\alpha B_2)\, A_1) = \alpha \operatorname{Tr}(\hat{S}(B_1 | B_2)\, A_1)$, para se ter o resultado. ∎

Recupera-se o Corolário 3.4.1 do Corolário 3.4.2 quando $\alpha = \beta = 1$.

Corolário 3.4.3 *Se $A_1, B_1 \geq 0$ e $A_2, B_2 > 0$, então a função*

$$f(\alpha) = \log \operatorname{Tr}\left((A_1 \#_\alpha A_2)(B_1 \#_\alpha B_2)\right), \qquad 0 \leq \alpha \leq 1,$$

satisfaz $f'(0) \leq f(1) - f(0)$.

Demonstração: Dividindo ambos os membros da desigualdade (3.31) por $\operatorname{Tr}(A_1 B_1)$, aplicando a exponencial e multiplicando ambos os membros da desigualdade assim obtida por $\operatorname{Tr}(A_1 B_1)$, podemos reescrever (3.31) na forma

$$\operatorname{Tr}(A_1 B_1) \exp \frac{\operatorname{Tr}(\hat{S}(A_1 | A_2)\, B_1) + \operatorname{Tr}(\hat{S}(B_1 | B_2)\, A_1)}{\operatorname{Tr}(A_1 B_1)} \leq \operatorname{Tr}(A_2 B_2). \tag{3.32}$$

Da definição de f, é claro que $\mathrm{e}^{f(0)} = \operatorname{Tr}(A_1 B_1)$ e $\mathrm{e}^{f(1)} = \operatorname{Tr}(A_2 B_2)$. Por outro lado, tomando a derivada de f em relação a α na origem e recordando (3.3), obtém-se

$$f'(0) = \frac{\operatorname{Tr}(\hat{S}(A_1 | A_2)\, B_1 + A_1\, \hat{S}(B_1 | B_2))}{\operatorname{Tr}(A_1 B_1)}.$$

Assim, podemos reescrever (3.32) na forma $\mathrm{e}^{f(0)} \mathrm{e}^{f'(0)} \leq \mathrm{e}^{f(1)}$ e tem-se o pretendido. ∎

Apresenta-se, de seguida, uma generalização da desigualdade de Peierls-Bogoliubov.

Corolário 3.4.4 *Dadas matrizes Hermíticas G, H, K e L, tem-se*

$$\operatorname{Tr}(\mathrm{e}^H \mathrm{e}^G) \exp \frac{\operatorname{Tr}(\hat{S}(\mathrm{e}^H | \, \mathrm{e}^{G+L})\, \mathrm{e}^G) + \operatorname{Tr}(\hat{S}(\mathrm{e}^G | \, \mathrm{e}^{H+K})\, \mathrm{e}^H)}{\operatorname{Tr}(\mathrm{e}^H \mathrm{e}^G)} \leq \operatorname{Tr}(\mathrm{e}^{H+K} \mathrm{e}^{G+L}). \tag{3.33}$$

Se $G = L = 0$, (3.33) *reduz-se à desigualdade de Peierls-Bogoliubov.*

Demonstração: A desigualdade (3.33) é consequência imediata de tomar $A_1 = \mathrm{e}^H$, $A_2 = \mathrm{e}^{G+L}$, $B_1 = \mathrm{e}^G$ e $B_2 = \mathrm{e}^{H+K}$ em (3.32). No caso particular $G = L = 0$, tem-se

$$\operatorname{Tr} \hat{S}(\mathrm{e}^H | \, I_n) = -\operatorname{Tr}(\mathrm{e}^H H), \qquad \operatorname{Tr}(\hat{S}(I_n | \, \mathrm{e}^{H+K})\, \mathrm{e}^H) = \operatorname{Tr}(\mathrm{e}^H (H+K)),$$

e (3.33) reduz-se à desigualdade de Peierls-Bogoliubov. ∎

3.5 Desigualdade Termodinâmica

Uma das primeiras desigualdades envolvendo o traço de matrizes aplicada à Mecânica Estatística é a *desigualdade de Golden-Thompson*. Em 1965, Sidney Golden [69], Colin J. Thompson [126] e Kurt Symanzik [124] provaram, independentemente, que

$$\mathrm{Tr}\left(\mathrm{e}^{H+K}\right) \;\leq\; \mathrm{Tr}\left(\mathrm{e}^{H}\mathrm{e}^{K}\right) \tag{3.34}$$

vale para H e K matrizes Hermíticas. Thompson [126] observou que a desigualdade (3.34) pode ser utilizada para obter um limite superior para a função de partição $Z = \mathrm{Tr}\,(\mathrm{e}^{-\hat{H}/\Theta})$ de uma cadeia antiferromagnética, dada uma partição adequada do Hamiltoniano $\hat{H}$. Por outro lado, Golden [69] obteve uma cadeia de limites inferiores para a *função energia-livre de Helmholtz*:

$$F \;=\; -\Theta \log \mathrm{Tr}\,(\mathrm{e}^{-\hat{H}/\Theta}) \tag{3.35}$$

de um sistema físico de Hamiltoniano $\hat{H}$. Desde então, esta desigualdade tem sido generalizada e complementada de vários modos (vide [2, 77, 79] e suas referências).

Desigualdades Tipo Golden-Thompson

Na Biologia da População e Teoria de Controle [31, 44], o comportamento de certos modelos matemáticos depende de outras funções dos valores próprios de e^{H+K} e $\mathrm{e}^{H}\mathrm{e}^{K}$, com H, K matrizes Hermíticas, para além do traço.

Dada $A \in M_n$ e $n_k = \binom{n}{k}$, se os valores próprios da k-ésima matriz composta de A estão ordenados por ordem não-crescente de valor absoluto:

$$\left|\lambda_1(A^{(k)})\right| \;\geq\; \cdots \;\geq\; \left|\lambda_{n_k}(A^{(k)})\right|,$$

define-se o *traço parcial* de A por

$$\mathrm{Tr}_m^{(k)}(A) \;=\; \sum_{i=1}^{m} \lambda_i(A^{(k)}), \qquad k = 1, \ldots, n, \quad m = 1, \ldots, n_k. \tag{3.36}$$

Por exemplo, se $m = n_k$, então (3.36) é a k-ésima função simétrica elementar dos valores próprios de A. Em particular, $\mathrm{Tr}_n^{(1)}(A) = \mathrm{Tr}(A)$ e $\mathrm{Tr}_1^{(n)}(A) = \det(A)$. Se $A \geq 0$ e $m = 1$, então (3.36) reduz-se, pela propriedade C3, ao produto dos k maiores valores próprios de A. Assim, dadas $A, B \in M_n$, desigualdades entre traços parciais do tipo

$$\mathrm{Tr}_m^{(k)}(A) \;\leq\; \mathrm{Tr}_m^{(k)}(B), \qquad k = 1, \ldots, n, \quad m = 1, \ldots, n_k,$$

incluem, em particular:

- a majoração fraca $A \prec_w B$ (quando $k = 1$);
- desigualdades entre as k-ésimas funções simétricas elementares dos valores próprios (quando $m = n_k$), entre as quais $\det(A) \leq \det(B)$;
- a majoração-log $A \prec_{(\log)} B$ (quando $k = n$), se $A, B \geq 0$ têm igual determinante.

No Teorema 3.5.1, estende-se ao traço parcial a variante mais forte da desigualdade de Golden-Thompson de Fumio Hiai e Dénes Petz [77, Teorema 1.1].

Teorema 3.5.1 *Se $H, K \in M_n$ são matrizes Hermíticas e $p > 0$, então*

$$\mathrm{Tr}_m^{(k)}(\mathrm{e}^{H+K}) \;\leq\; \mathrm{Tr}_m^{(k)}\big(\mathrm{e}^{pH/2}\mathrm{e}^{pK}\mathrm{e}^{pH/2}\big)^{1/p}, \qquad k = 1, \ldots, n, \quad m = 1, \ldots, n_k. \tag{3.37}$$

e o segundo membro de (3.37) *converge decrescentemente para o primeiro membro, à medida que $p \downarrow 0$.*

Demonstração: Atendendo à majoração-log (3.14), com $A = (\mathrm{e}^H)^{(k)}$ e $B = (\mathrm{e}^K)^{(k)}$, e às propriedades C1-C2 das matrizes compostas, tem-se

$$\Big(\big(\mathrm{e}^{qH/2}\mathrm{e}^{qK}\mathrm{e}^{qH/2}\big)^{1/q}\Big)^{(k)} \;\prec_{(\log)}\; \Big(\big(\mathrm{e}^{pH/2}\mathrm{e}^{pK}\mathrm{e}^{pH/2}\big)^{1/p}\Big)^{(k)}, \qquad 0 < q \leq p, \tag{3.38}$$

para $k = 1, \ldots, n$. Tomando limites quando $q \downarrow 0$, pela versão paramétrica da fórmula de Lie-Trotter, e dado que a majoração logarítmica (3.38), envolvendo matrizes $n_k \times n_k$, implica a majoração fraca, tem-se

$$\big(\mathrm{e}^{H+K}\big)^{(k)} \;\prec_w\; \Big(\big(\mathrm{e}^{pH/2}\mathrm{e}^{pK}\mathrm{e}^{pH/2}\big)^{1/p}\Big)^{(k)}, \qquad p > 0, \qquad k = 1, \ldots, n,$$

o que, em termos de traços parciais, corresponde às desigualdades (3.37). ■

Nota 3.5.1 Claramente, se $k = p = 1$ e $m = n$, pela propriedade cíclica do traço, (3.37) reduz-se à desigualdade de Golden-Thompson.

Decompondo convenientemente o Hamiltoniano numa soma $\hat{H} = H + K$ (em particular, pode encarar-se H como a energia cinética e K como a energia potencial), considera-se

$$F_p \;=\; -\Theta \log \mathrm{Tr}\left(\mathrm{e}^{-pH/(2\Theta)}\mathrm{e}^{-pK/\Theta}\mathrm{e}^{-pH/(2\Theta)}\right)^{1/p}, \qquad p > 0.$$

Do Teorema 3.5.1, com $k = 1$ e $m = n$, pela monotonia crescente da função logaritmo, tem-se $F_p \leq F_q \leq F$, $0 < q \leq p$. Num modelo clássico, H e K comutam, a função de partição coincide com $Z_1 = \mathrm{Tr}(\mathrm{e}^{-H/\Theta}\mathrm{e}^{-K/\Theta})$ e $F_1 = -\Theta \log \mathrm{Tr}\,(Z_1)$ corresponde a uma estimativa, dita *pseudo-clássica*, da função energia-livre de Helmholtz. Esta comutatividade de H e K não ocorre num modelo quântico. Assim, de $F_1 \leq F$, conclui-se que a função de Helmholtz clássica fornece uma aproximação, um limite inferior, para a função de Helmholtz quântica correcta.

A desigualdade de Golden-Thompson foi ainda complementada por Fumio Hiai e Dénes Petz [77] do seguinte modo:

$$\mathrm{Tr}\left(\mathrm{e}^{pH}\#_\alpha \mathrm{e}^{pL}\right)^{1/p} \;\leq\; \mathrm{Tr}\left(\mathrm{e}^{(1-\alpha)H+\alpha L}\right), \qquad p>0, \tag{3.39}$$

quaisquer que sejam as matrizes Hermíticas H, L e $0 \leq \alpha \leq 1$.

Desigualdade Termodinâmica

Os observáveis quânticos são modelados por matrizes Hermíticas. Por exemplo, o operador energia H é um operador Hermítico. O *valor médio estatístico da energia* do estado descrito pela matriz densidade A é igual a $E = \mathrm{Tr}\,(AH)$ e a *energia livre* desse estado é dada pela função

$$\psi(A) \;=\; \mathrm{Tr}\,(AH) - \Theta\, S(A), \tag{3.40}$$

Determinar um extremo da função $\psi(A)$ é um problema importante, em Física. Por conveniência, considera-se $\Theta = -1$.

Teorema 3.5.2 *Se H é uma matriz Hermítica, então*

$$\log \mathrm{Tr}\,(\mathrm{e}^H) \;=\; \max\left\{\mathrm{Tr}\,(AH) + S(A):\;\; A \geq 0,\;\; \mathrm{Tr}(A) = 1\right\}.$$

Demonstração: Seja H uma matriz Hermítica fixa. Pretendemos maximizar a função $\psi(A) = \mathrm{Tr}\,(AH) + S(A)$ no conjunto das matrizes densidade. Seja S uma matriz Hermítica e considere-se a função diferenciável $f(t) = \psi\,(\mathrm{e}^{-itS}A\mathrm{e}^{itS})$, para $t \in \mathbb{R}$ numa vizinhança da origem. Atendendo a que e^{itS} é uma matriz unitária e relembrando que a entropia é invariante por transformações de semelhança unitária, tem-se

$$f(t) \;=\; \mathrm{Tr}\left(\mathrm{e}^{-itS}A\mathrm{e}^{itS}H\right) + S(A),$$

cuja derivada na origem é igual a $f'(0) = i\,\mathrm{Tr}\,(S[H,A])$. Pretendemos determinar as matrizes densidade A para as quais f atinge um extremo em $t = 0$. Da condição de extremo $f'(0) = 0$ e dada a arbitrariedade da matriz S, conclui-se que $[H,A] = 0$, o que implica que $[\mathrm{e}^H, A] = 0$. Atendendo a que $\mathrm{Tr}\,(A) = 1$, tem-se

$$\begin{aligned}
\psi(A) \;&=\; \mathrm{Tr}\left(A\big(\log \mathrm{e}^H - \log A\big)\right)\\
&=\; \log \mathrm{Tr}\,(\mathrm{e}^H) - \mathrm{Tr}\left(A\big(\log \mathrm{Tr}\,(\mathrm{e}^H) + \log A - \log \mathrm{e}^H\big)\right)\\
&=\; \log \mathrm{Tr}\,(\mathrm{e}^H) - \mathrm{Tr}\left(\mathrm{e}^H A\,\mathrm{e}^{-H} \log\left(\mathrm{Tr}\,(\mathrm{e}^H) A\,\mathrm{e}^{-H}\right)\right),
\end{aligned}$$

sendo possível escrever a última igualdade dada a comutabilidade entre as matrizes e^H e A. Tomando $B = \mathrm{Tr}\,(e^H)A\,e^{-H}$ na anterior expressão, vem

$$\begin{aligned}\psi(A) &= \log \mathrm{Tr}\,(e^H) - \mathrm{Tr}\left(e^H B \log B\right)/\mathrm{Tr}(e^H) \\ &= \log \mathrm{Tr}\,(e^H) - \mathrm{Tr}\left(e^H(B\log B - B + I)\right)/\,\mathrm{Tr}\,(e^H),\end{aligned}$$

já que $\mathrm{Tr}\,(e^H B) = \mathrm{Tr}\,(e^H)$. Observa-se, ainda, que $[e^H, B\log B - B + I] = 0$. Atendendo a que $x\log x - x + 1 \geq 0$, $x \geq 0$, com igualdade se e só se $x = 1$, as matrizes e^H e $B\log B - B + I$, que são simultaneamente diagonalizáveis, têm por produto uma matriz semi-definida positiva. Portanto, conclui-se que

$$\psi(A) \;\leq\; \log \mathrm{Tr}\,(e^H), \tag{3.41}$$

como pretendido. Além disso, a ocorrência de igualdade em (3.41) implica que B seja a matriz identidade, donde $A = e^H/\,\mathrm{Tr}\,(e^H)$. Por outro lado, se $A = e^H/\,\mathrm{Tr}\,(e^H)$, então A é uma matriz densidade que satisfaz $\psi(A) = \log \mathrm{Tr}\,(e^H)$. ■

O Teorema 3.5.2 estabelece a importante *desigualdade termodinâmica* [82]:

$$\log \mathrm{Tr}\,(e^H) \;\geq\; \mathrm{Tr}\,(AH) + S(A), \tag{3.42}$$

que vale para A uma matriz densidade e H uma matriz Hermítica. Além disso, ocorre igualdade em (3.42) se e só se $A = e^H/\,\mathrm{Tr}\,(e^H)$. Substituindo H por $-H/\Theta$ na desigualdade termodinâmica e multiplicando ambos os membros pelo simétrico de Θ, constata-se que a função $\psi(A)$ apresentada em (3.40) funciona como uma aproximação, limite superior, para a função energia livre de Helmholtz F definida em (3.35).

Em particular, se H é a matriz nula no Teorema 3.5.2, confirma-se que $\log n$ é a entropia máxima $S(A)$, no conjunto das matrizes densidade, e que esta é atingida para $A = I_n/n$. Por outro lado, substituindo H por $H + \log B$ no Teorema 3.5.2, tem-se o seguinte resultado de Fumio Hiai e Dénes Petz [77].

Corolário 3.5.1 *Se $B > 0$ e H é uma matriz Hermítica, então*

$$\log \mathrm{Tr}\,(e^{H+\log B}) \;=\; \max\left\{\mathrm{Tr}\,(AH) - S(A,B) :\; A \geq 0,\; \mathrm{Tr}(A) = 1\right\}.$$

É óbvio que o Teorema 3.5.2 é um caso particular do Corolário 3.5.1, quando $B = I_n$.

Generalizamos, de seguida, a desigualdade termodinâmica, utilizando a variante mais forte da desigualdade de Golden-Thompson, ou seja, o caso tracial do Teorema 3.5.1.

Teorema 3.5.3 *Sejam $A > 0, B > 0$, com $\mathrm{Tr}(A) = 1$ e H Hermítica. Então*

$$\log \mathrm{Tr}\left(\mathrm{e}^{qH/2}(A^p \#_\alpha B^p)^{q/p}\, \mathrm{e}^{qH/2}\right)^{1/q} \;\geq\; \mathrm{Tr}\,(AH) + \frac{\alpha}{p}\,\mathrm{Tr}\,(\hat{S}\,(A^p|B^p)\,A^{1-p}), \tag{3.43}$$

para cada $0 \leq \alpha \leq 1$ e $p, q > 0$. Se $\alpha = q = 1$ e $B = I_n$, então (3.43) reduz-se à desigualdade termodinâmica.

Demonstração: Se $A > 0$ e $B > 0$, então $(A^p \#_\alpha B^p)^{1/p} > 0$, para qualquer $0 \leq \alpha \leq 1$ e $p > 0$. Atendendo à variante mais forte da desigualdade de Golden-Thompson, aplicada às matrizes Hermíticas H e $K = \log\,(A^p \#_\alpha B^p)^{1/p}$, e à monotonia da função logaritmo, tem-se

$$\log \mathrm{Tr}\left(\mathrm{e}^{qH/2}(A^p \#_\alpha B^p)^{q/p}\, \mathrm{e}^{qH/2}\right)^{1/q} \;\geq\; \log \mathrm{Tr}\left(\mathrm{e}^{H + \log(A^p \#_\alpha B^p)^{1/p}}\right), \quad 0 \leq \alpha \leq 1, \;\; p, q > 0.$$

Substituindo a matriz B por $(A^p \#_\alpha B^p)^{1/p}$ no Corolário 3.5.1, verifica-se o seguinte:

$$\log \mathrm{Tr}\left(\mathrm{e}^{H + \log(A^p \#_\alpha B^p)^{1/p}}\right) \;\geq\; \mathrm{Tr}\,(AH) - S(A, (A^p \#_\alpha B^p)^{1/p}), \qquad 0 \leq \alpha \leq 1, \;\; p > 0.$$

As anteriores desigualdades em conjunto com a desigualdade tracial logarítmica (3.26) permitem obter (3.43). Para concluir a demonstração toma-se $\alpha = q = 1$. Obtém-se

$$\log \mathrm{Tr}\,(\mathrm{e}^H B) \;\geq\; \mathrm{Tr}\,(AH) + \frac{1}{p}\,\mathrm{Tr}\,(\hat{S}\,(A^p|B^p)\,A^{1-p}), \qquad p > 0.$$

Além disso, se $B = I_n$, então $\mathrm{Tr}\,(\hat{S}\,(A^p|B^p)\,A^{1-p}) = p\,S(A)$ e tem-se a desigualdade termodinâmica. ■

3.6 Cadeia de Equivalências entre Desigualdades

A seguinte cadeia de desigualdades inclui, em particular, a versão generalizada da desigualdade termodinâmica apresentada no Teorema 3.5.3, quando $q = 1$.

Teorema 3.6.1 *As seguintes condições valem e são equivalentes, para cada $p > 0$ e $0 \leq \alpha \leq 1$:*

(i) *Sejam $A > 0$, $B > 0$, com $\mathrm{Tr}(A) = 1$ e D Hermítica. Então*

$$\log \mathrm{Tr}\left(\mathrm{e}^D (A^p \#_\alpha B^p)^{1/p}\right) \;\geq\; \mathrm{Tr}\,(AD) + \frac{\alpha}{p}\,\mathrm{Tr}\,(\hat{S}(A^p|B^p)A^{1-p}).$$

(ii) *Dadas matrizes Hermíticas H, K e L, tem-se*

$$\mathrm{Tr}\,(\mathrm{e}^H)\,\exp\frac{p\,\mathrm{Tr}\,(\mathrm{e}^H(H+K)) + \alpha\,\mathrm{Tr}\,(\hat{S}(\mathrm{e}^{pH}|\mathrm{e}^{pL})\,\mathrm{e}^{(1-p)H})}{p\,\mathrm{Tr}\,(\mathrm{e}^H)} \;\leq\; \mathrm{Tr}\left(\mathrm{e}^{H+K}(\mathrm{e}^{pH} \#_\alpha \mathrm{e}^{pL})^{1/p}\right).$$

(iii) *Se* $A_1 > 0, A_2 > 0$ *e* $B_2 > 0$, *então*

$$\frac{1}{n} S\left(\mathrm{Tr}\,(A_1)\, I_n, \mathrm{Tr}\big((A_1^p \#_\alpha A_2^p)^{1/p} B_2\big)\, I_n\right) \;\leq\; -\frac{\alpha}{p}\,\mathrm{Tr}\,(\hat{S}(A_1^p|A_2^p)\, A_1^{1-p}) - \mathrm{Tr}\,(\hat{S}(I_n|B_2)\, A_1).$$

Demonstração (i) $\Rightarrow$ (iii) : Dadas $A_1, A_2, B_2 > 0$, as matrizes $A = A_1/\,\mathrm{Tr}\,(A_1)$, $B = A_2$ e $D = \log B_2$ estão nas condições da alínea (i) e satisfazem as propriedades seguintes:

$$\begin{aligned} (A^p \#_\alpha B^p)^{1/p} &= (\mathrm{Tr}\,(A_1))^{\alpha-1}\,(A_1^p \#_\alpha A_2^p)^{1/p}, && (3.44)\\ \mathrm{Tr}\,(AD) &= \frac{\mathrm{Tr}\,(\hat{S}(I_n|B_2)\, A_1)}{\mathrm{Tr}\,(A_1)}, && (3.45)\\ \mathrm{Tr}\,(\hat{S}(A^p|B^p)\, A^{1-p}) &= \frac{\mathrm{Tr}\,(\hat{S}(A_1^p|A_2^p) A_1^{1-p})}{\mathrm{Tr}\,(A_1)} + p \log \mathrm{Tr}\,(A_1). && (3.46)\end{aligned}$$

Substituindo (3.44), (3.45) e (3.46) na desigualdade em (i) e multiplicando ambos os membros pelo simétrico de $\mathrm{Tr}\,(A_1)$, tem-se

$$\begin{aligned} &-\mathrm{Tr}\,(A_1)\left((\alpha-1)\log \mathrm{Tr}\,(A_1) + \log \mathrm{Tr}\left(B_2 (A_1^p \#_\alpha A_2^p)^{1/p}\right)\right) \leq && (3.47)\\ &\leq -\mathrm{Tr}\,(\hat{S}(I_n|B_2)\, A_1) - \frac{\alpha}{p}\,\mathrm{Tr}\,(\hat{S}(A_1^p|A_2^p)\, A_1^{1-p}) - \alpha\, \mathrm{Tr}\,(A_1) \log \mathrm{Tr}\,(A_1). \end{aligned}$$

Como a soma do primeiro membro de (3.47) com $\alpha\, \mathrm{Tr}(A_1) \log \mathrm{Tr}(A_1)$ é igual a

$$\frac{1}{n} S\left(\mathrm{Tr}\,(A_1)\, I_n, \mathrm{Tr}\left((A_1^p \#_\alpha A_2^p)^{1/p} B_2\right) I_n\right),$$

prova-se a implicação pretendida.

(iii) $\Rightarrow$ (ii) : Sejam H, K e L matrizes Hermíticas. Considerando $A_1 = \mathrm{e}^H$, $A_2 = \mathrm{e}^L$ e $B_2 = \mathrm{e}^{H+K}$, tem-se

$$\begin{aligned} \frac{1}{n} S\big(\mathrm{Tr}\,(A_1) I_n, \mathrm{Tr}\big((A_1^p \#_\alpha A_2^p)^{1/p} B_2\big)\, I_n\big) &= -\mathrm{Tr}\,(\mathrm{e}^H) \log \frac{\mathrm{Tr}\big(\mathrm{e}^{H+K}(\mathrm{e}^{pH} \#_\alpha \mathrm{e}^{pL})^{1/p}\big)}{\mathrm{Tr}\,(\mathrm{e}^H)}, && (3.48)\\ \mathrm{Tr}\,(\hat{S}(I_n|B_2) A_1) &= \mathrm{Tr}\,(\mathrm{e}^H (H+K)). && (3.49)\end{aligned}$$

Substituindo (3.48) e (3.49) na desigualdade em (iii) e dividindo ambos os membros da desigualdade assim obtida por $-\mathrm{Tr}\,(\mathrm{e}^H)$, tem-se

$$\log \frac{\mathrm{Tr}\left(\mathrm{e}^{H+K}(\mathrm{e}^{pH} \#_\alpha \mathrm{e}^{pL})^{1/p}\right)}{\mathrm{Tr}\,(\mathrm{e}^H)} \;\geq\; \frac{\alpha\, \mathrm{Tr}\,(\hat{S}(\mathrm{e}^{pH}|\mathrm{e}^{pL})\, \mathrm{e}^{(1-p)H})}{p\, \mathrm{Tr}\,(\mathrm{e}^H)} + \frac{\mathrm{Tr}\,(\mathrm{e}^H (H+K))}{\mathrm{Tr}\,(\mathrm{e}^H)}.$$

Aplicando a função exponencial, vem

$$\frac{\mathrm{Tr}\left(\mathrm{e}^{H+K}(\mathrm{e}^{pH} \#_\alpha \mathrm{e}^{pL})^{1/p}\right)}{\mathrm{Tr}\,(\mathrm{e}^H)} \;\geq\; \exp \frac{p\, \mathrm{Tr}\,(\mathrm{e}^H(H+K)) + \alpha\, \mathrm{Tr}\,(\hat{S}(\mathrm{e}^{pH}|\mathrm{e}^{pL})\, \mathrm{e}^{(1-p)H})}{p\, \mathrm{Tr}\,(\mathrm{e}^H)}.$$

(ii) $\Rightarrow$ (i) : Se $A, B > 0$, então existem matrizes Hermíticas H, L tais que $A = \mathrm{e}^H$ e $B = \mathrm{e}^L$. Sendo H Hermítica, então $K = D - \log A$ também é Hermítica e (ii) implica que

$$\mathrm{Tr}\,(A) \exp \frac{p\,\mathrm{Tr}\,(AD) + \alpha \mathrm{Tr}\,(\hat{S}(A^p|B^p)A^{(1-p)})}{p\,\mathrm{Tr}\,(A)} \;\leq\; \mathrm{Tr}\left(\mathrm{e}^D(A^p\#_\alpha B^p)^{1/p}\right).$$

Por hipótese, $\mathrm{Tr}\,(A) = 1$ e pela monotonia da função logaritmo, obtém-se

$$\mathrm{Tr}\,(AD) + \frac{\alpha}{p}\,\mathrm{Tr}\,(\hat{S}(A^p|B^p)A^{1-p}) \;\leq\; \log \mathrm{Tr}\left(\mathrm{e}^D(A^p\#_\alpha B^p)^{1/p}\right). \qquad \blacksquare$$

Como constataremos, existe uma interessante relação entre as desigualdades da anterior cadeia de equivalências, quando $p = 1$, e o Corolário 3.4.2, com $B_1 = I_n$.

Corolário 3.6.1 *As seguintes condições verificam-se e são equivalentes, qualquer que seja* $0 \leq \alpha \leq 1$*:*

(i) *Sejam* $A > 0, B > 0$, *com* $\mathrm{Tr}\,(A) = 1$ *e* D *Hermítica. Então*

$$\log \mathrm{Tr}\left(\mathrm{e}^D(A\#_\alpha B)\right) \;\geq\; \mathrm{Tr}\,(AD) + \alpha \mathrm{Tr}\,(\hat{S}(A|B)).$$

(ii) *Dadas matrizes Hermíticas* H, K *e* L, *tem-se*

$$\mathrm{Tr}\,(\mathrm{e}^H)\,\exp\,\frac{\mathrm{Tr}\,(\mathrm{e}^H(H+K)) + \alpha\,\mathrm{Tr}\,\hat{S}(\mathrm{e}^H|\mathrm{e}^L)}{\mathrm{Tr}\,(\mathrm{e}^H)} \;\leq\; \mathrm{Tr}\left(\mathrm{e}^{H+K}(\mathrm{e}^H\#_\alpha \mathrm{e}^L)\right).$$

(iii) *Se* $A_1 > 0$, $A_2 > 0$ *e* $B_2 > 0$, *então*

$$\frac{1}{n} S\,(\mathrm{Tr}\,(A_1)\,I_n, \mathrm{Tr}\,((A_1\#_\alpha A_2)B_2)\,I_n) \;\leq\; -\alpha\,\mathrm{Tr}\,(\hat{S}(A_1|A_2)) - \mathrm{Tr}\,(\hat{S}(I_n|B_2)\,A_1).$$

Demonstração: Basta tomar $p = 1$ no Teorema 3.6.1. $\blacksquare$

Observa-se, facilmente, que o Corolário 3.6.1 (iii), substituindo a matriz B_2 por B_2^β, $0 \leq \beta \leq 1$, se reduz ao Corolário 3.4.2, com $B_1 = I_n$.

No caso particular em que $\alpha = 1$ e para uma escolha adequada das matrizes envolvidas, a cadeia de equivalências antes estabelecida envolve desigualdades bem conhecidas. Se $B = I_n$, conforme já referido, o Teorema 3.6.1 (i) contempla a desigualdade termodinâmica. Por outro lado, se $L = 0$, então

$$\left(\mathrm{e}^{pH}\#_1 \mathrm{e}^{pL}\right)^{1/p} \;=\; I_n, \qquad \mathrm{Tr}\,(\hat{S}(\mathrm{e}^{pH}|\mathrm{e}^{pL})\mathrm{e}^{(1-p)H}) \;=\; -p\,\mathrm{Tr}\,(\mathrm{e}^H H)$$

e o Teorema 3.6.1 (ii) reduz-se à desigualdade de Peierls-Bogoliubov. Em resumo, obtém-se o seguinte importante resultado, que estabelece a equivalência entre a desigualdade termodinâmica, a desigualdade de Peierls-Bogoliubov e a positividade da entropia relativa para matrizes densidade.

Corolário 3.6.2 *As seguintes condições verificam-se e são equivalentes:*

(i) **(Desigualdade Termodinâmica)** *Seja $A > 0$, com $\mathrm{Tr}\,(A) = 1$ e D Hermítica. Então*

$$\log \mathrm{Tr}(\mathrm{e}^D) \;\geq\; \mathrm{Tr}(AD) + S(A),$$

ocorrendo igualdade se e só se $A = \mathrm{e}^D / \mathrm{Tr}\,(\mathrm{e}^D)$;

(ii) **(Desigualdade de Peierls-Bogoliubov)** *Dadas matrizes Hermíticas H e K, tem--se*

$$\mathrm{Tr}\,(\mathrm{e}^H)\,\exp \frac{\mathrm{Tr}\,(\mathrm{e}^H K)}{\mathrm{Tr}\,(\mathrm{e}^H)} \;\leq\; \mathrm{Tr}\,(\mathrm{e}^{H+K}),$$

ocorrendo igualdade se e só se K é uma matriz escalar;

(iii) *Se $A_1 > 0$ e $A_2 > 0$, então*

$$\frac{1}{n} S\left(\mathrm{Tr}\,(A_1)\, I_n, \mathrm{Tr}\,(A_2)\, I_n\right) \;\leq\; S(A_1, A_2),$$

ocorrendo igualdade se e só se $\mathrm{Tr}\,(A_2)\, A_1 = \mathrm{Tr}\,(A_1)\, A_2$;

(iv) *Se $A > 0$ e $B > 0$ são matrizes densidade, então*

$$S(A, B) \;\geq\; 0,$$

ocorrendo igualdade se e só se $A = B$.

Demonstração: As equivalências entre as afirmações (i), (ii) e (iii) resultam do Corolário 3.6.1, com $\alpha = 1$, sendo fácil verificar as ocorrências de igualdade.

(iii) $\Rightarrow$ (iv) : Pode reescrever-se a desigualdade em (iii) na forma abreviada:

$$S\left(A_1 / \mathrm{Tr}\,(A_1), A_2 / \mathrm{Tr}\,(A_2)\right) \;\geq\; 0. \tag{3.50}$$

Substituindo A_1, A_2 em (3.50) pelas matrizes A, B de traço um, tem-se $S(A, B) \geq 0$. A caracterização da igualdade em (iii) garante que $S(A, B) = 0$ se e só se $A = B$.

(iv) $\Rightarrow$ (i) : Aplicando (iv) às matrizes densidade $A > 0$ e $B = \mathrm{e}^D / \mathrm{Tr}\,(\mathrm{e}^D)$, com D Hermítica, tem-se $S(A, \mathrm{e}^D / \mathrm{Tr}\,(\mathrm{e}^D)) \geq 0$. Da definição de entropia relativa, atendendo a que e^D comuta com a matriz escalar $\mathrm{Tr}\,(\mathrm{e}^D)\, I_n$, vem

$$\begin{aligned} S(A, \mathrm{e}^D / \mathrm{Tr}\,(\mathrm{e}^D)) &= \mathrm{Tr}\left(A(\log A - \log \mathrm{e}^D + \log\left(\mathrm{Tr}\,(\mathrm{e}^D)\, I_n\right)\right) \\ &= -S(A) \;-\; \mathrm{Tr}\,(AD) \;+\; \log \mathrm{Tr}\,(\mathrm{e}^D), \end{aligned}$$

cuja positividade mostra a validade da desigualdade termodinâmica. ■

O Corolário 3.6.2 (iii) fornece um limite inferior para a entropia relativa $S(A_1, A_2)$, que apenas depende dos traços das matrizes $A_1, A_2 > 0$ envolvidas.

Para terminar esta secção, tecem-se algumas considerações, envolvendo algumas das desigualdades anteriormente obtidas. Dadas H e L matrizes Hermíticas, $0 \leq \alpha \leq 1$ e $p > 0$, consideram-se as fórmulas:

$$\begin{aligned} l_{\alpha,p}(H,L) &= \operatorname{Tr}(\mathrm{e}^H) \exp \frac{\alpha \operatorname{Tr}(\hat{S}(\mathrm{e}^{pH}|\mathrm{e}^{pL})\, \mathrm{e}^{(1-p)H})}{p \operatorname{Tr}(\mathrm{e}^H)}, \\ u_\alpha(H,L) &= \operatorname{Tr}(\mathrm{e}^{(1-\alpha)H+\alpha L}). \end{aligned}$$

Como consequência de (3.21), com $A = \mathrm{e}^H$ e $B = \mathrm{e}^L$, e atendendo à desigualdade de Peierls-Bogoliubov, com $K = \alpha(L - H)$, obtém-se

$$l_{\alpha,p}(H,L) \;\leq\; \operatorname{Tr}(\mathrm{e}^H) \exp \frac{\alpha \operatorname{Tr}(\mathrm{e}^H(L-H))}{\operatorname{Tr}(\mathrm{e}^H)} \;\leq\; u_\alpha(H,L). \tag{3.51}$$

Por outro lado, tomando $K = -H$ no Teorema 3.6.1 (ii) e pela desigualdade de Golden--Thompson complementada (3.39), tem-se

$$l_{\alpha,p}(H,L) \;\leq\; \operatorname{Tr}(\mathrm{e}^{pH} \#_\alpha \mathrm{e}^{pL})^{1/p} \;\leq\; u_\alpha(H,L). \tag{3.52}$$

Assim, comparar as duas expressões em (3.51) e (3.52), que admitem o mesmos limites inferior $l_{\alpha,p}(H,L)$ e superior $u_\alpha(H,L)$, torna-se uma questão naturalmente interessante.

Ora, para cada $0 < \alpha_0 \leq 1$ e H, L matrizes Hermíticas, tais que $L - H$ é uma matriz não escalar, existe $p_0 > 0$ (dependente de H, L, α_0) para o qual vale

$$\operatorname{Tr}(\mathrm{e}^H) \exp \frac{\alpha \operatorname{Tr}(\mathrm{e}^H(L-H))}{\operatorname{Tr}(\mathrm{e}^H)} \;<\; \operatorname{Tr}(\mathrm{e}^{pH} \#_\alpha \mathrm{e}^{pL})^{1/p}, \tag{3.53}$$

para todo o $0 < p \leq p_0$ e $\alpha_0 \leq \alpha \leq 1$. De facto, dadas matrizes Hermíticas H, L tais que $L - H$ é uma matriz não escalar, suponhamos, para qualquer $\epsilon > 0$, que existe p, tal que $0 < p \leq \epsilon$, e

$$\operatorname{Tr}(\mathrm{e}^H) \exp \frac{\alpha \operatorname{Tr}(\mathrm{e}^H(L-H))}{\operatorname{Tr}(\mathrm{e}^H)} \;\geq\; \operatorname{Tr}(\mathrm{e}^{pH} \#_\alpha \mathrm{e}^{pL})^{1/p}. \tag{3.54}$$

Seja $\mathcal{P} = \{p > 0 : (3.54) \text{ vale}\}$. Este conjunto tem zero como ponto de acumulação. Considerando uma sucessão em $\mathcal{P}$ convergente para a origem, tomando limites à medida que $p \downarrow 0$ nessa sucessão e recordando a fórmula (3.22), a partir de (3.54), tem-se

$$\operatorname{Tr}(\mathrm{e}^H) \exp \frac{\alpha \operatorname{Tr}(\mathrm{e}^H(L-H))}{\operatorname{Tr}(\mathrm{e}^H)} \;\geq\; \operatorname{Tr}\left(\mathrm{e}^{(1-\alpha)H+\alpha L}\right). \tag{3.55}$$

Mas a desigualdade de Peierls-Bogoliubov apenas permite a ordem inversa no sinal da desigualdade anterior, concluindo-se que só a igualdade pode ocorrer em (3.55). A condição de igualdade na desigualdade de Peierls-Bogoliubov implica que $L - H$ é uma matriz escalar, o que contradiz a hipótese. Assim se prova a existência de $p_0 > 0$, tal que (3.53) se verifica. Além disso, se $L - H = \lambda I_n$, então reduzem-se ambos os membros de (3.53) a $e^{\alpha\lambda}\operatorname{Tr}(e^H)$, ocorrendo a igualdade em (3.53), para qualquer $p > 0$.

Acabámos de confirmar que, para cada $0 < \alpha_0 \leq 1$, existe $p_0 > 0$, tal que

$$l_{\alpha,p}(H, L) \;\leq\; \operatorname{Tr}(e^H)\exp\frac{\alpha\operatorname{Tr}(e^H(L-H))}{\operatorname{Tr}(e^H)} \;\leq\; \operatorname{Tr}(e^{pH}\#_\alpha e^{pL})^{1/p} \;\leq\; u_\alpha(H, L),$$

para todo o $0 < p \leq p_0$ e $\alpha_0 \leq \alpha \leq 1$, ocorrendo igualdade, quando $L - H$ é uma matriz escalar.

Considerações Finais

Não encontro significado para a palavra fim. Quando olho para um círculo, não distingo onde ele começa e onde ele acaba.

Filipe Lascasas

No final desta Dissertação, cumpre fazer um balanço do trabalho desenvolvido. O ponto de partida, que motivou o título desta Dissertação, foi o estudo do contradomínio numérico de alguns operadores e de algumas desigualdades matriciais, com aplicações em Física. Fomos, deste modo, conduzidos aos conteúdos originais dos segundo e terceiro capítulos.

Muito embora, os obstáculos encontrados apenas tivessem permitido, numa fase preliminar, obter resultados parciais para o contradomínio numérico de operadores de emparelhamento bosónicos (agora completamente caracterizados no capítulo dois), são estes mesmos obstáculos que nos levam a questionar a utilidade de investigar o contradomínio numérico-J para matrizes A de ordem dois. Assim, da necessidade de completar esta discussão, surgiu o nosso interesse pelos contradomínios numéricos em espaços de Krein. Aliás, a organização dos conteúdos desta Dissertação segue um lógica um pouco diversa da ordem cronológica pela qual foram obtidos.

Compreende-se, por isso, o realce dado ao Teorema do Contradomínio Hiperbólico no capítulo inicial. Este fornece, a par com o Teorema do Contradomínio Elíptico, a descrição completa dos contradomínios numéricos em espaços de Krein de matrizes de dimensão dois. A distinção entre os seus diversos casos degenerados, agora óbvia e bem demarcada, era completamente desconhecida, no início da nossa investigação. Estes foram-se clarificando à medida que o estudo paralelo do contradomínio numérico clássico do operador de emparelhamento se fazia. Este fora afinal o mote de partida. Esta interligação sugeriu, por exemplo, a formulação final do resultado, por alíneas dependentes do sinal de duas constantes associadas à matriz A e aos seus valores próprios.

A passagem para o contradomínio tracial-C, J e para o contradomínio determinantal-C, J surgiu, logo depois, na medida em que o Teorema do Contradomínio Hiperbólico permitiu derivar, de forma quase imediata e natural, as caracterizações destes conjuntos para matrizes de ordem dois, quando C é diagonal, e ainda as suas formas especiais de conjunto singular e de subconjunto de uma recta.

Terminamos, com alguns problemas e questões em aberto que consideramos de interesse nesta linha de investigação.

Problemas em Aberto

Somewhere, there is something incredible waiting to be known.

Carl Sagan

Várias classes de matrizes admitem contradomínios numéricos elípticos, independentemente da sua ordem. É o caso de algumas matrizes tridiagonais [36, 39, 40, 41] e das matrizes quadráticas [128]. As matrizes tridiagonais unitariamente semelhantes a

$$A = \begin{bmatrix} aI_r & X \\ Y & bI_{n-r} \end{bmatrix}, \qquad a, b \in \mathbb{C}, \quad X, Y^* \in M_{r\times(n-r)}, \tag{3.56}$$

em que XY e YX são matrizes normais, $p = \min(r, n-r)$, $0 < r < n$, têm por contradomínio numérico o invólucro convexo de p elipses [36]. (Na Secção 1.2.5, investigámos o contradomínio numérico-J, $J = I_r \oplus -I_{n-r}$, de tal classe de matrizes por blocos.) A seguinte questão é pertinente.

Problema 1: Estudar em que condições o contradomínio numérico-c, $c \in \mathbb{R}^n$, de uma matriz tridiagonal unitariamente semelhante a A do tipo (3.56) é um disco elíptico ou o invólucro convexo de elipses, e descrevê-lo.

Utilizando a caracterização paramétrica (Lema 2.4.1) da curva geradora de fronteira de $W_c(A)$, $c \in \mathbb{R}^n$, obtemos uma elipse, no caso particular em que a matriz tridiagonal, de elemento genérico a_{ij}, tem diagonal principal constante e as entradas fora da diagonal principal satisfazem a seguinte condição:

$$\exists\, k \in \mathbb{C}: \qquad a_{jj+1} = k\,\overline{a_{j+1j}}, \quad j \in J_1, \qquad \text{e} \qquad a_{jj+1} = k\,\overline{a_{j+1j}}, \quad j \in J_2,$$

onde J_1, J_2 formam uma partição de $\{1, \ldots, n\}$. Podemos, assim, atacar o Problema 1, sempre que $a = b$ e $Y = kX^*$ na matriz (3.56). (Alguns avanços estão em curso [27]). Contudo, a situação mais geral, $a \neq b$, com XY e YX matrizes normais, está ainda por investigar.

A partir do momento em que nos propusemos levar os contradomínios numéricos a navegar às águas menos conhecidas dos produtos internos indefinidos, alguns outros problemas se nos colocaram.

Problema 2: Se $H \in H_2$ não-singular é indefinida, caracterizar completamente os conjuntos $W_C^H(A)$ e $\triangle_C^H(A)$, quaisquer que sejam $A \in M_2$ e $C \in M_2$.

No primeiro capítulo desta Dissertação, referimos os Teoremas do Contradomínio Elíptico para os conjuntos $W_C(A)$ e $\triangle_C(A)$, sendo $C \in M_2$ diagonal. Estes resultados foram recentemente estendidos a $C \in M_2$ arbitrária. Hiroshi Nakazato [107] demonstrou que $W_C(A)$ tem a fronteira elíptica, quaisquer que sejam $A, C \in M_2$ (e Chi-Kwong Li [94] simplificou a demonstração). Natália Bebiano e Graça Soares [24] mostraram a forma elíptica de $\triangle_C(A)$, para $A, C \in M_2$ arbitrárias. Portanto, se $H \in H_2$ é definida positiva, atendendo às relações (1.48) e (1.62), facilmente se obtêm as formas geométricas de $W_C^H(A)$ e $\triangle_C^H(A)$ a partir das já conhecidas para $W_C(A)$ e $\triangle_C(A)$, respectivamente. Permanece em aberto o caso indefinido. Os Teoremas 1.3.2 e 1.4.2 apenas fornecem a solução do Problema 2, quando $C \in M_2$ é diagonal, $H = \operatorname{diag}(1, -1)$ e $A \in M_2$ é arbitrária.

Problema 3: Dadas $J = I_r \oplus -I_{n-r}$, $0 < r < n$, $A \in M_n$ Hermítica-J e $C \in M_n(\mathbb{R})$, encontrar uma condição suficiente (necessária e suficiente) para $W_C^J(A)$ ser todo o eixo real, em termos dos valores próprios de A.

Problema 4: Dada $J = I_r \oplus -I_{n-r}$, $0 < r < n$, caracterizar o contradomínio tracial-C, J de $A \in M_n$, uma matriz Hermítica-J, quando $C \in M_n(\mathbb{R})$ é diagonal e todos os valores próprios de A são reais.

O Corolário 1.3.1 apresenta a solução dos Problemas 3 e 4, quando $n = 2$. Alguns resultados parciais estão em desenvolvimento, para o caso em que a matriz $A \in M_n$ possui r valores próprios reais em $\sigma_J^+(A)$ e $n - r$ valores próprios reais em $\sigma_J^-(A)$ [26]. Se algum valor próprio de A admitir vectores próprios isotrópicos-J, a questão configura-se de maior complexidade.

O problema da representação paramétrica da fronteira de $W_H(A)$ merece, também, a nossa atenção. Uma hipótese a considerar seria seguir as técnicas de Mao-Ting Chien e Hiroshi Nakazato [42] na parametrização da curva geradora de fronteira do conjunto $W_c(A)$, $c \in \mathbb{R}^n$, e a teoria de polinómios hiperbólicos de Gårding.

Problema 5: Dadas $A, C \in M_n$ e $H \in H_n$ não-singular, provar que as fronteiras de $W_C^H(A)$ e $\triangle_C^H(A)$ são uniões finitas de arcos algébricos e, portanto, curvas de classe C^∞, excepto num número finito de pontos, seguindo uma abordagem alternativa à considerada na Secção 1.3.2.

Seria interessante investigar, por exemplo, se a abordagem seguida por H. Nakazato, Y. Nishikawa e M. Takaguchi [108] para provar que $\partial W_C(A)$ é uma união finitas de arcos algébricos admite um paralelo para o conjunto $W_C^H(A)$.

Quando $C \in M_n$, o contradomínio numérico-C de $A \in M_n$ não é, em geral, convexo. Mas, se $C \in M_n(\mathbb{R})$ é uma matriz diagonal, o Teorema de Westwick confirma que $W_C(A)$ é um conjunto convexo, qualquer que seja $A \in M_n$. A questão que se segue é naturalmente relevante.

Problema 6: Se $J = I_r \oplus -I_{n-r}$, $0 < r < n$, e $C \in M_n(\mathbb{R})$ é diagonal, mostrar a convexidade ou não do contradomínio tracial-C, J de $A \in M_n$.

Se o Problema 6 admitir uma resposta afirmativa, então criam-se novos caminhos de investigação, nomeadamente, a possibilidade de gerar computacionalmente um traçado aproximado para a curva geradora de fronteira do conjunto $W_C^J(A)$. O Problema 6 afigura--se, contudo, de elevado grau de dificuldade.

Um famoso resultado variacional para matrizes Hermíticas é o *Princípio do Máximo de Ky Fan* (1950) [103] que estabelece que

$$\max_{\Lambda} \sum_{i=1}^{k} \langle Ax_i, x_i \rangle \; = \; \sum_{i=1}^{k} \alpha_i, \qquad k = 1, \ldots, n,$$

onde Λ é um conjunto $\{x_1, \ldots, x_k\}$ de k vectores ortonormados em $\mathbb{C}^n$ e $\alpha_1 \geq \cdots \geq \alpha_n$ são os valores próprios da matriz Hermítica $A \in M_n$. Podemos provar este célebre resultado, utilizando a teoria dos contradomínios numéricos, atendendo ao facto de os extremos do contradomínio numérico-k de uma matriz $A \in M_n$, $k = 1, \ldots, n$, serem dados por pontos-σ do conjunto.

O Princípio do Máximo de Ky Fan é uma fonte de inspiração, muitas vezes, usado na prova de outros resultados. Dele facilmente se deriva o *Teorema de Schur* (1923) [103], o primeiro exemplo de majoração na história da análise matricial, que relaciona os valores próprios e as entradas da diagonal principal de uma matriz Hermítica.

Problema 7: Dada $J = I_r \oplus -I_{n-r}$, $0 \leq r \leq n$, averiguar em que condições se podem estender à classe das matrizes Hermíticas-J algumas desigualdades clássicas para matrizes Hermíticas, como o Princípio do Máximo de Ky Fan e o Teorema de Schur.

O conhecido *Teorema de Birkhoff* (1946) [103], válido para matrizes duplamente estocásticas, afirma que o conjunto destas matrizes constitui o invólucro convexo das matrizes de permutação. No estudo dos contradomínios numéricos em espaços de Krein, somos naturalmente conduzidos à consideração de matrizes estocásticas-J. A título exemplificativo, se $n = 3$ e $J = \mathrm{diag}(1, 1, -1)$, uma matriz $A = [a_{ij}] \in M_3(\mathbb{R})$ denomina-se *duplamente estocástica-J* se a soma das entradas de cada linha e de cada coluna de A é igual à unidade e se são satisfeitas as seguintes condições:

$$\begin{array}{llll} a_{ij} \geq 0, & i, j = 1, 2 & \text{ou} & i = j = 3; \\ a_{ij} \leq 0, & i = 1, 2,\ j = 3 & \text{ou} & i = 3,\ j = 1, 2. \end{array}$$

Problema 8: Estender o Teorema de Birkhoff a matrizes duplamente estocásticas-J.

Muito recentemente, Tsuyoshi Ando [4] estabeleceu a desigualdade de Löwner de tipo indefinido, no âmbito de matrizes A, B Hermíticas-J de valores próprios reais não negativos e da relação de ordem $A \geq^J B$ definida por $\langle Ax, x\rangle_J \geq \langle Bx, x\rangle_J$, $x \in \mathbb{C}^n$. Neste contexto, Ando mostrou que

$$I \geq^J A \geq^J B \quad \Rightarrow \quad f(A) \geq^J f(B),$$

qualquer que seja a função matricial monótona f definida num intervalo real, contendo os valores próprios de A e B, obtendo, em particular, a versão indefinida da desigualdade de Löwner-Heinz, quando $f(t) = t^{1/2}$, $t \in [0, +\infty)$. Observa-se que o espectro de uma matriz A Hermítica-J satisfazendo $I \geq^J A$ é real. Como vimos na Secção 3.2, a desigualdade de Furuta generaliza a desigualdade de Löwner-Heinz clássica.

Problema 9: Investigar a possibilidade de obter uma versão de tipo indefinido para a desigualdade de Furuta.

Ainda no âmbito dos estudos desenvolvidos nos dois últimos capítulos, colocam-se outras questões de maior nível de dificuldade, entre as quais:

Problema 10: Investigar propriedades espectrais de operadores de emparelhamento definidos na álgebra de Grassmann ou na álgebra simétrica sobre $\mathbb{C}^n$, $n \geq 3$.

Problema 11: Estudar o contradomínio numérico da primeira (ou da segunda) derivação de matrizes de ordem n superior (ou igual) a dois, em $\mathbb{C}^n_{(m)}$, $m \in \mathbb{N}$.

Problema 12: Utilizar a técnica da majoração logarítmica na obtenção de novas desigualdades traciais logarítmicas.

Apresentámos alguns exemplos de questões que desafiam a nossa curiosidade e que merecem empenhada atenção. Importa, por isso, prosseguir à descoberta de novos rumos, que permitam alargar o alcance do horizonte do saber actual.

Bibliografia

The library is the mathematician's laboratory.

P. R. Halmos, in *I Want to Be a Mathematician*

[1] T. Ando, *Topics on Operator Inequalities*, Lecture Notes, Hokkaido Univ., Sapporo, 1978.

[2] T. Ando e F. Hiai, Log-majorization and complementary Golden-Thompson type inequalities, *Linear Algebra Appl.* **197** (1994), 113-131.

[3] T. Ando e F. Hiai, Hölder type inequalities for matrices, *Math. Ineq. Appl.* **1** (1998), 1-30.

[4] T. Ando, Löwner inequality of indefinite type, *Linear Algebra Appl.* **385** (2004), 73-80.

[5] H. Araki, On an inequality of Lieb and Thirring, *Lett. Math. Phys.* **19** (1990), 167--170.

[6] Y. H. Au-Yeung e N.-K. Tsing, An extension of the Hausdorff-Toeplitz theorem on the numerical range, *Proc. Amer. Math. Soc.* **89** (1983), 215-218.

[7] R. Baierlein, How entropy got its name, *Amer. J. Phys.* **60** (1992), 1151.

[8] N. Bebiano, *Contradomínios Numéricos Generalizados de Matrizes - Variações Sobre este Tema*, Dissertação de Doutoramento, Universidade de Coimbra, 1984.

[9] N. Bebiano e J. Queiró, The determinant of the sum of two normal matrices with prescribed eigenvalues, *Linear Algebra Appl.* **71** (1985), 23-28.

[10] N. Bebiano, Some variations on the concept of the C-numerical range, *Portugal. Math.* **43** (1985/6), 189-200.

[11] N. Bebiano, Nondifferentiable points of $\partial W_c(A)$, *Linear Multilin. Algebra* **19** (1986), 249-257.

[12] N. Bebiano, Some analogies between the c-numerical range and a certain variation on this concept, *Linear Algebra Appl.* **81** (1986), 47-54.

[13] N. Bebiano, J. K. Merikoski e J. da Providência, On a conjecture of G. N. de Oliveira on determinants, *Linear Multilin. Algebra*, **20** (1987), 167-170.

[14] N. Bebiano, Y. Poon e J. da Providência, On C-det-spectral and C-det-convex matrices, *Linear Multilin. Algebra* **23** (1988), 343-351.

[15] N. Bebiano, C.-K. Li e J. da Providência, Some results on the numerical range of a derivation, *SIAM J. Matrix Anal. Appl.* **14** (1993), 1084-1095.

[16] N. Bebiano, New developments on the Marcus-Oliveira conjecture, *Linear Algebra Appl.*, **197/198** (1994), 793-803.

[17] N. Bebiano e J. da Providência, Numerical ranges in physics, *Linear Multilin. Algebra* **43** (1998), 327-337.

[18] N. Bebiano, C.-K. Li e J. da Providência, Generalized numerical ranges of permanental compounds arising from quantum systems of bosons, *Electr. J. Linear Algebra* **7** (2000), 73-91.

[19] N. Bebiano, R. Lemos, J. da Providência e G. Soares, Remarks on generalized numerical ranges of operators on an indefinite inner product space, *Online Proceedings of SIAM Applied Linear Algebra Conference*, College of William and Mary, Williamsburg, 15-19 de Julho, 2003.

[20] N. Bebiano, J. da Providência Jr. e R. Lemos, Matricial inequalities in statistical mechanics, *Linear Algebra Appl.* **376** (2004), 256-273.

[21] N. Bebiano, R. Lemos e J. da Providência, Numerical ranges of unbounded operators arising in quantum physics, *Linear Algebra Appl.* **381** (2004), 259-279.

[22] N. Bebiano, R. Lemos, J. da Providência e G. Soares, On generalized numerical ranges of operators on an indefinite inner product space, *Linear Multilin. Algebra* **52** (2004), 203-233.

[23] N. Bebiano, R. Lemos e J. da Providência, Inequalities for quantum relative entropy, *Linear Algebra Appl.*, em publicação (no volume dedicado a G. N. de Oliveira).

[24] N. Bebiano e G. Soares, Three observations on the determinantal range, *Linear Algebra Appl.*, em publicação (no volume dedicado a G. N. de Oliveira).

[25] N. Bebiano, R. Lemos, J. da Providência e G. Soares, On the geometry of numerical ranges in spaces with an indefinite inner product, *Linear Algebra Appl.*, em publicação.

[26] N. Bebiano, H. Nakazato, J. da Providência, R. Lemos e G. Soares, Inequalities for J-Hermitian matrices, submetido.

[27] N. Bebiano, R. Lemos, J. da Providência e G. Soares, Matrices with elliptical c--numerical range, em preparação.

[28] V. P. Belavkin e P. Staszewski, C^*-algebraic generalization of relative entropy and entropy, *Ann. Inst. H. Poincaré* Sect. A **37** (1982), 51-58.

[29] F. A. Berezin, *The Method of Second Quantization*, Academic Press, New York, 1966.

[30] C. A. Berger, *Normal Dilations*, Ph.D. Thesis, Cornell University, 1963.

[31] D. S. Bernstein, Inequalities for the trace of matrix exponentials, *SIAM J. Matrix Anal. Appl.* **9** (1988), 156-158.

[32] R. Bhatia, *Matrix Analysis*, Springer-Verlag, New York, 1997.

[33] J. P. Blaizot e G. Ripka, *Quantum Theory of Finite Systems*, MIT Press, Cambridge, 1986.

[34] F. Bonsall e J. Duncan, Numerical ranges, in: R. G. Bartle (ed), *Studies in functional analysis* **21**, Math. Assoc. America, Washington, 1980, 1-49.

[35] E. Brieskorn e H. Knörrer, *Plane Algebraic Curves*, Birkhäuser Verlag, Basel, 1986.

[36] E. Brown e I. Spitkovsky, On matrices with elliptical numerical ranges, *Linear Multilin. Algebra* **52** (2004), 177-193.

[37] N. N. Chan e M. K. Kwong, Hermitian matrices and a conjecture, *Amer. Math. Monthly* **92** (1985), 533-541.

[38] W.-S. Cheung e N.-K. Tsing, The C-numerical range of matrices is star-shaped, *Linear Multilin. Algebra* **41** (1996), 245-250.

[39] M.-T. Chien, On the numerical range of tridiagonal operators, *Linear Algebra Appl.*, **246** (1996), 203-214.

[40] M.-T. Chien, The envelope of the generalized numerical range, *Linear Multilin. Algebra*, **43** (1998), 363-376.

[41] M.-T. Chien, L. Yeh, Y. T. Yeh and F. Z. Lin, On geometric properties of the numerical range, *Linear Algebra Appl.*, **274** (1998), 389-410.

[42] M.-T. Chien e H. Nakazato, Boundary generating curves of the c-numerical range, *Linear Algebra Appl.*, **294** (1999), 67-84.

[43] M.-T. Chien e H. Nakazato, The c-numerical range of tridiagonal matrices, *Linear Algebra Appl.*, **335** (2001), 55-61.

[44] J. E. Cohen, S. Friedland, T. Kato e F. P. Kelly, Eigenvalue inequalities for products of matrix exponentials, *Linear Algebra Appl.* **45** (1982), 55-95.

[45] J. B. Conway, *A Course in Functional Analysis*, Springer-Verlag, New York, 1985.

[46] Ch. Davis, The Toeplitz-Hausdorff theorem explained, *Canad. Math. Bull.* **14** (1971), 245-246.

[47] W. F. Donoghue, On the numerical range of a bounded operator, *Mich. Math. J.* **4** (1957), 261-263.

[48] A. L. Duarte, *O Teorema de Tarski e suas Aplicações em Teoria de Matrizes*, Universidade de Coimbra, 1993.

[49] S. W. Drury e B. Cload, On the determinantal conjecture of Marcus and de Oliveira, *Linear Algebra Appl.* **177** (1992), 105-109.

[50] *Encyclopaedia of Mathematics, Supplement I*, Kluwer Academic Publishers, 1997.

[51] A. Feintuch e E. Markus, The Toeplitz-Hausdorff theorem and robust stability theory, *Math. Intellinger* **21** (1999), 33-35.

[52] M. Fiedler, Bounds for the determinant of the sum of Hermitian matrices, *Proc. Amer. Math. Soc.* **30** (1987), 27-31.

[53] M. Fiedler, Geometry of the numerical range of matrices, *Linear Algebra Appl.* **37** (1981), 81-96.

[54] J. I. Fujii e E. Kamei, Relative operator entropy in noncomutative information theory, *Math. Japon.* **34** (1989), 341-348.

[55] J. I. Fujii e T. Furuta, Löwner-Heinz, Cordes and Heinz-Kato inequalities, *Math. Japon.* **38** (1993), 73-78.

[56] W. Fulton, *Algebraic Curves*, Benjamin, Massachusetts, 1964.

[57] S. Furuichi, K. Yanagi e K. Kuriyama, Fundamental properties of Tsallis relative entropy, *J. Math. Phys.* **45** (2004), 4868-4877.

[58] T. Furuta, $A \geq B \geq 0$ ensures $(B^r A^p B^r)^{1/q} \geq B^{(p+2r)/q}$, for $r \geq 0, p \geq 0, q \geq 1$ with $(1+2r)q \geq p+2r$, *Proc. Amer. Math. Soc.* **101** (1987), 85-88.

[59] T. Furuta, An elementary proof of an order preserving inequality, *Proc. Japan Acad.* **65A** (1989), 126.

[60] T. Furuta, A proof via operator mean of an order preserving inequality, *Linear Algebra Appl.* **113** (1989), 129-130.

[61] T. Furuta, Norm inequalities equivalent to Löwner-Heinz theorem, *Reviews in Math. Phys.* **1** (1989), 135-137.

[62] T. Furuta, An extension of the Furuta inequality and Ando-Hiai log-majorization, *Linear Algebra Appl.* **219** (1994), 139-155.

[63] T. Furuta, *Invitation to Linear Operators - From Matrices to Bounded Linear Operators on a Hilbert space*, Taylor and Francis, London, 2001.

[64] T. Furuta, Convergence of logarithmic trace inequalities via generalized Lie-Trotter Formulae, *Linear Algebra Appl.* **396** (2005), 353-372.

[65] T. Furuta, Inequalities associated with Umegaki relative entropy $S(A,B) = \mathrm{Tr}(A \log A - A \log B)$ and the relative operator entropy $\hat{S}(A|B) = A^{\frac{1}{2}}(A^{-\frac{1}{2}} B A^{-\frac{1}{2}}) A^{\frac{1}{2}}$ by J. I. Fujii and E. Kamei, pré-publicação.

[66] I. Gohberg, P. Lancaster e L. Rodman, *Matrices and Indefinite Scalar Products*, Birhäuser Verlag, Basel, 1983.

[67] M. Goldberg e E. G. Straus, Elementary inclusion relations for generalized numerical ranges, *Linear Algebra Appl.* **18** (1977), 1-24.

[68] M. Goldberg, On certain finite dimensional ranges and numerical radii, *Linear Multlin. Algebra* **7** (1979), 329-342.

[69] S. Golden, Lower bounds for the Helmholtz function, *Phys. Rev.* **137** (1965), B1127--B1128.

[70] R. D. Grigorieff e R. Platto, On the minimax equality for seminorms, *Linear Algebra Appl.* **221** (1995), 227-243.

[71] K. Gustafson e D. Rao, *Numerical Range - The Field of Values of Linear Operators and Matrices*, Springer, New York, 1997.

[72] E. Gutkin, E. A. Jonckheer e M. Karow, Convexity of the joint numerical range: topological and differential geometric viewpoint, *Linear Algebra Appl.* **376** (2004), 143-171.

[73] P. R. Halmos, *A Hilbert Space Problem Book*, Van Nostrand, New York, 1967.

[74] F. Hansen e G. K. Pedersen, Jensen's inequality for operators and Löwner's theorem, *Math. Ann.* **258** (1982), 229-241.

[75] F. Hausdorff, Der Wertvorrat einer Bilinearform, *Math. Z.* **3** (1919), 314-316.

[76] E. Heinz, Beiträge sur Störungstheorie der Spectralzerlegung, *Math. Ann.* **123** (1951), 415-438.

[77] F. Hiai e D. Petz, The Golden-Thompson trace inequality is complemented, *Linear Algebra Appl.* **181** (1993), 153-185.

[78] F. Hiai, Equality cases in matrix norm inequalities of Golden-Thompson type, *Linear Multilin. Algebra* **35** (1994), 239-249.

[79] F. Hiai, Log-majorizations and norm inequalities for exponential operators, in: *Linear Operators*, Banach Center Publ. 38, Inst. Math., Polish Acad. Sci., Warszawa, 1997, 119-181.

[80] R. A. Horn e C. R. Johnson, *Matrix Analysis*, Cambridge Univ. Press, New York, 1985.

[81] R. A. Horn e C. R. Johnson, *Topics in Matrix Analysis*, Cambridge Univ. Press, New York, 1991.

[82] K. Huang, *Statistical Mechanics*, John Wiley and Sons, New York, 1987.

[83] N. Jacobson, *Basic Algebra I*, Second edition, W. H. Freeman and Company, New York, 1985.

[84] T. Kato, Notes on some inequalities for linear operators, *Math. Ann.* **125** (1952), 208-212.

[85] D. Keeler, L. Rodman e I. Spitkovsky, The numerical range of 3×3 matrices, *Linear Algebra Appl.* **252** (1997), 115-139.

[86] R. Kippenhahn, Über den Wertvorrat einer Matrix, *Math. Nach.* **6** (1951), 193-228.

[87] F. Kubo e T. Ando, Means of positive linear operators, *Math. Ann.* **246** (1980), 205-224.

[88] S. Kullback e R. A. Leibler, On information and sufficiency, *Ann. Math. Statist.* **22** (1951), 79-86.

[89] P. Lancaster and M. Tismenetsky, *The Theory of Matrices: Second Edition with Applications*, Academic Press, New York, 1985.

[90] L. D. Landau e E. M. Lifshitz, *Statistical Physics*, Adison-Wesley Publishing Company, Massachusetts, 1958.

[91] R. Lemos, *Desigualdades Espectrais e Majoração Logarítmica*, Dissertação de Mestrado, Universidade de Coimbra, 1998.

[92] C.-K. Li, C-Numerical Ranges and C-Numerical Radii, *Linear Multilin. Algebra* **37** (1994), 51-82.

[93] C.-K. Li, A simple proof of the elliptical range theorem, *Proc. Amer. Math. Soc.* **124**, 7 (1996), 1985-1986.

[94] C.-K. Li, Some convexity theorems for the generalized numerical ranges, *Linear Multilin. Algebra* **40** (1996), 235-240.

[95] C.-K. Li, N.-K. Tsing e F. Uhlig, Numerical ranges of an operator in an indefinite inner product space, *Electr. J. Linear Algebra* **1** (1996), 1-17.

[96] C.-K. Li e L. Rodman, Shapes and computer generation of numerical ranges of Krein space operators, *Electr. J. Linear Algebra* **3** (1998), 31-47.

[97] C.-K. Li e L. Rodman, Remarks on numerical ranges of operators in spaces with an indefinite metric, *Proc. Amer. Math. Soc.* **126** (1998), 973-982.

[98] E. Lieb e W. Thirring, in: E. Lieb, B. Simon e A. S. Wightman (eds), *Studies in Mathematical Physics*, Princeton Univ. Press, Princeton, 1976, 301-302.

[99] K. Löwner, Über monotone Matrixfunctionen, *Math. Z.* **38** (1934), 177-216.

[100] M. Marcus, Derivations, Plücker relations and numerical range, *Indiana Univ. Math. J.* **22** (1973), 1137-1149.

[101] M. Marcus, *Finite Dimensional Multilinear Algebra I*, Marcel Dekker, New York, 1973.

[102] M. Marcus e I. Filippenko, Nondifferentiable boundary points of higher numerical range, *Linear Algebra Appl.* **21** (1978), 217-232.

[103] A. W. Marshall e I. Olkin, *Inequalities: Theory of Majorization and its Applications*, Academic Press, New York, 1979.

[104] M. E. Miranda, *Sobre Valores Singulares de Matrizes Complexas*, Dissertação de Doutoramento, Universidade de Coimbra, 1982.

[105] F. D. Murnaghan, On the field of values of a square matrix, *Proc. Nat. Acad. Sci.* **18** (1932), 246-248.

[106] M. Nakamura e H. Umegaki, A note on entropy for operator algebras, *Proc. Jap. Acad.* **37** (1961), 149-154.

[107] H. Nakazato, The C-numerical range of a 2×2 matrix, *Sci. Rep. Hirozaki Univ.* **41** (1994), 197-206.

[108] H. Nakazato, Y. Nishikawa e M. Takaguchi, On the boundary of the C-numerical range of a matrix, *Linear Multilin. Algebra* **39** (1995), 231-240.

[109] H. Nakazato e K. Tsumura, k-Numerical range and the structural performance of buildings, *Sci. Math. Japonicae* **53** (2000), 101-117.

[110] J. W. Negele e H. Orland, *Quantum Many-Particle Systems*, Addison-Wesley Publishing Company, New York, 1988.

[111] J. von Neumann, Thermodynamik quantummechanischer Gesamheiten, *Gött. Nachr.* **1** (1927), 273-291.

[112] M. Ohya e D. Petz, *Quantum Entropy and Its Use*, Springer-Verlag, Berlin, 1993.

[113] G. N. de Oliveira, Normal matrices (research problem), *Linear Multilin. Algebra* **12** (1982), 153-154.

[114] A. L. Onishchik, *Lie Groups and Lie Algebras*, Springer-Verlag, Berlin, 1993.

[115] G. K. Pedersen, Some operator monotone functions, *Proc. Amer. Math. Soc.* **36** (1972), 309-310.

[116] Y.-T. Poon, Another proof of a result of Westwick, *Linear Multilin. Algebra* **9**, (1980), 35-37.

[117] J. da Providência, The numerical ranges of derivations and quantum physics, *Linear Multilin. Algebra* **37** (1994), 213-220.

[118] P. J. Psarrakos e M. J. Tsatsomeros, *Numerical Range: (in) a Nutshell*, Math. Notes, Washington State Univ., Vol 45, N.º 2, 2002 e Vol 46, N.º 1, 2003.

[119] M. Reed e B. Simon, *Methods of Modern Mathematical Physics II*, Academic Press, New York, 1966.

[120] C. E. Shannon, A mathematical theory of communication, *Bell System Tech. Jour.* **27** (1948), 379-423 e 623-656.

[121] H. Shapiro, A conjecture of Kippenhahn about the characteristic polynomial of a pencil generated by two Hermitian matrices II, *Linear Algebra Appl.* **45** (1982), 97-108.

[122] J. H. Shapiro, *Notes on the Numerical Range*, Lecture Notes, Michigan State Univ., 2004.

[123] M. H. Stone, *Linear Transformations in Hilbert Spaces*, Americ. Math. Soc., Providence, RI, 1932.

[124] K. Symanzik, Proofs and refinements of an inequality of Feynman, *J. Math. Phys.* **6** (1965), 1155-1156.

[125] K. Tanahashi, Best possibility of Furuta inequality, *Proc. Amer. Math. Soc.* **128** (1996), 141-146.

[126] C. J. Thompson, Inequality with applications in statistical mechanics, *J. Math. Phys.* **6** (1965), 1812-1813.

[127] O. Toeplitz, Das algebraische Analogon zu einem Satze von Fejér, *Math. Z.* **2** (1918), 187-197.

[128] S.-H. Tso e P. Y. Wu, Matricial ranges of quadratic operators, *Rocky Mountain J. Math.* **29** (1999), 1139-1152.

[129] H. Umegaki, Condition expectation in an operator algebra IV, *Kodai Math. Sem. Rep.* **14** (1962), 59-85.

[130] R. J. Walker, *Algebraic Curves*, Princeton Univ. Press, Princeton, 1950.

[131] R. Westwick, A theorem on numerical range, *Linear Multilin. Algebra* **2** (1975), 311-315.

[132] A. Wintner, Sur Theorie der beschränkten Bilinearformen, *Math. Z.* **30** (1929), 228-282.

[133] K. Yanagi, K. Kuriyama e S. Furuichi, Generalized Shannon inequalities based on Tsallis relative operator entropy, *Linear Algebra Appl.* **394** (2005), 109-118.